

Realisierung eines Wellenfrontsensors mit einem ASIC

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Universität Mannheim

vorgelegt von
Dipl.-Ing. Dirk Droste
aus Erlangen

Mannheim, 1999

Dekan:	Prof. Dr. Reinhard Männer, Universität Mannheim
Referent:	Prof. Dr. Reinhard Männer, Universität Mannheim
Korreferent:	Prof. Dr. Josef Bille, Universität Heidelberg

Tag der mündlichen Prüfung: 8. Juli 1999

Zum Andenken an meinen Vater

Zusammenfassung

In dieser Arbeit werden die Entwicklung und der Test eines applikationsspezifischen integrierten Schaltkreises (ASIC) und der Aufbau eines Hartmann-Shack-Wellenfrontsensors unter Anwendung des ASIC beschrieben. Der integrierte Schaltkreis ermöglicht durch eine Kombination von Photodetektoren und Signalverarbeitungselementen die für das Hartmann-Shack-Verfahren erforderliche Mustererkennung einer Matrix aus Fokalpunkten. Die Anwendung des Hartmann-Shack-Wellenfrontsensors in einem ophthalmologischen Visualisierungsverfahren zur Diagnose von Retina-Erkrankungen, um die optischen Aberrationen des zu untersuchenden Auges zu quantifizieren, definiert zusammen mit der Systematik des Hartmann-Shack-Verfahrens die Spezifikationen zur Entwicklung des ASIC.

Ein wesentlicher Teil der Arbeit bestand in der Entwicklung der Detektionseinheit des ASIC zur Bestimmung der Position der Fokalpunkte. Die ophthalmologische Anwendung definiert eine Wiederholrate der Positionsdetektion von 1 kHz bei einer Strahlungsleistung der Fokalpunkte von 1 nW. Um die zur Rekonstruktion der Phaseninformation erforderliche Genauigkeit der Positionsdetektion zu ermöglichen, wurde in dem ASIC eine analoge Signalverarbeitung implementiert, die über die Bestimmung des stärksten Photostromes unter mehreren eine Quantisierung der Fokalpunkte-Positionen vornimmt. Die Positionsdaten werden anschließend an ein Datenverarbeitungssystem übertragen, um mit Hilfe eines Least-Square-Fit-Verfahrens die Koeffizienten einer Polynom-Funktion zur Beschreibung der optischen Aberrationen zu berechnen.

Das entwickelte und aufgebaute Sensor-System konnte unter den definierten Strahlungsleistungen der Fokalpunkte in einer Integrationszeit von 1 ms die optischen Aberration von Testlinsen mit einer absoluten Genauigkeit von ± 0.15 dpt vermessen, entsprechend einem 15% Fehler innerhalb des dynamischen Bereichs.

Abstract

In this thesis the development and the test of an application specific integrated circuit (ASIC) and the assembly of a Hartmann-Shack-Wavefrontdetector by use of the ASIC is described. The ASIC combines an array of photodetectors and modules of analog signal processing circuitry to perform the pattern recognition of a focal point pattern, generated by the Hartmann-Shack's lens array. The ophthalmologic application of the developed system to measure the eye's optical aberrations within a retina imaging system in combination with the Hartmann-Shack-Wavefrontdetection scheme defines the specification of the ASIC's performance.

A crucial part of this work was dedicated to the development of a position sensing module within the ASIC, which allows the measurements of the focal spot positions with a repetition rate of up to 1 kHz under illumination conditions of 1 nW incident power of each focal spot. To realize the specified accuracy of position detection, a winner-take-all circuit has been evaluated, which samples the spot positions for discrete values. The values are then transferred to a data processing unit to perform a least-square-fit of the generated values to calculate the coefficients of a standard polynomial function for wavefront description.

A system of a Hartmann-Shack-Wavefrontdetector has been set up, which allows measurements of optical aberrations of test-lenses under illumination conditions as mentioned above within a integration time of 1 ms and an absolute accuracy of ± 0.15 dpt according to an error of 15% of total dynamic range.

Inhaltsverzeichnis

Einleitung	1
1 Verbesserung der Retina-Mikromorphometrie	3
1.1 Anwendung von Retina-Scan-Tomographen	3
1.2 Die aberrationsbegrenzte Auflösung eines RST	4
1.3 Auflösungsverbesserung eines RST durch aktive Optik	6
1.4 Mathematische Beschreibung einer Wellenfront	7
1.4.1 Die Zernike-Polynome	8
1.5 Bestimmung einer Wellenfront	11
1.5.1 Das Hartmann-Shack-Detektionsverfahren	12
1.5.2 Die Rekonstruktion der Wellenfront	14
1.6 Anforderungen an den Wellenfrontsensor-ASIC	14
1.6.1 Auflösung und dynamischer Bereich der Detektion	15
1.6.2 Diskriminierungsgenauigkeit der Verschiebungsdetektion	17
1.6.3 Zeitliche Dauer der Detektion	18
1.6.4 Photosensitivität der Detektion	19
1.6.5 Zusammenfassung der Spezifikationen	19
2 Die optoelektronischen Möglichkeiten eines ASIC	21
2.1 Die Prozesstechnologien	22
2.2 Photodetektoren	23
2.2.1 Strahlungsabsorption	23
2.2.2 Photodioden	25
2.2.3 Phototransistoren	28
2.2.4 Gegenüberstellung der Photodetektoren	30
2.3 Der MOS-FET	31
2.3.1 Der Subthreshold-Betrieb	31
3 Die Entwicklung des ASIC	37
3.1 Systemübersicht	37
3.2 Der HSSX	38
3.2.1 Das Timing des HSSX	40
3.2.2 Das Pixel	40
3.2.3 Das Winner-Take-All Modul	40
3.2.4 Die dynamischen Schieberegister	58
3.2.5 Der Takt-Treiberbaum der dynamischen Schieberegister	63

3.2.6	Das Layout der WTA- und Schieberegister-Zelle	65
3.2.7	Der Kompressor	67
3.2.8	Die Steuerung des HSSX	73
3.2.9	Die Ein- und Ausgabe-Pads des HSSX	74
3.2.10	Das Layout des HSSX	75
3.3	Der Prototyp 1	76
3.4	Der Prototyp 2	79
3.4.1	Auflösungsverbesserung durch Interpolation	80
3.4.2	Die Signalverarbeitung des Cluster	81
3.4.3	Das WTA-Modul	82
3.4.4	Die Analogspeicher und der Ladungsverstärker	83
3.4.5	Das Digitalisierungsverfahren	84
3.5	Erkenntnisse aus der Entwicklung von Prototyp 1 und 2	86
4	Das externe System	89
4.1	Systemarchitektur	89
4.2	Die Schaltung des FPGA	91
4.3	Performance der Systeme	98
5	Anwendung des ersten Prototyps als Eye-Tracker	99
5.1	Photorefraktive Keratektomie	99
5.2	Detektion durch das Hartmann-Shack-Verfahren	100
5.3	Aufbau des Eye-Tracker	103
5.4	Software des Eye-Tracker	105
5.5	Messungen der Augenbewegungen eines Probanden	106
6	Messergebnisse des HSSX	109
6.1	Parameter des 0.6- μ m-CMOS-Prozess	109
6.2	Messung der Homogenität der Cluster-Matrix	117
6.3	Messung der Homogenität eines Cluster	118
6.4	Quantifizierung des Positionsrauschens	124
6.5	Messung der optischen Aberrationen von Test-Linsen	128
6.6	Leistungsaufnahme des HSSX	132
6.7	Diskussion der Messergebnisse	133
	Zusammenfassung	135
	A Least-Square-Fit	137
	B Verwendete Variablen	141
	Literaturverzeichnis	143

Einleitung

Die Glaukom-Früherkennung ist ein in der Augenheilkunde etabliertes Verfahren, um den Fortschritt einer typischen Form der Erkrankung des Auges zu bestimmen. Durch die fehlenden Abflussmöglichkeiten des Kammerwassers des Glaukom-betroffenen Auges entsteht ein erhöhter Innendruck und ein Absterben der Rezeptoren der Netzhaut ist die Folge.

Die in der Ophthalmologie eingesetzten Diagnose-Verfahren zur Erkennung der Glaukom-Erkrankung verwenden ein Visualisierungssystem, durch das die Struktur der Retina unter Abtasten mit einem Laserstrahl dreidimensional rekonstruiert werden kann. Durch die Darstellung lässt sich die topographische Gestalt der Retina in dem Bereich des Sehnervens bewerten und gegebenenfalls die pathologische Veränderung aufgrund einer Glaukom-Erkrankung diagnostizieren. Das zu untersuchende Auge selbst reduziert dabei die physikalisch erreichbare optische Auflösung des Verfahrens, da das System aus Linse und Hornhaut Aberrationen in den Messstrahl einbringt, die den Fokuspunkt des auf die Netzhaut fokussierten Laserstrahls aufweiten. Durch Kenntnis von Art und Höhe der optischen Aberrationen ist es möglich, durch ein steuerbares optisches Element den Messstrahl derart zu verformen, dass die Aberrationen kompensiert und der Fokuspunkt in seiner Größe verringert werden kann.

Die Quantifizierung der optischen Aberrationen des Auges kann mit Hilfe des Hartmann-Shack Wellenfront-Detektionsverfahrens erfolgen. In diesem Verfahren wird der auf Aberrationen zu untersuchende Lichtstrahl durch ein Linsenarray geführt, so dass eine charakteristische Matrix von Fokuspunkten entsteht. Die ortsabhängige Phaseninformation des Lichtstrahls wird aus der Position der einzelnen Fokuspunkte rekonstruiert und durch die Koeffizienten eines Polynoms höherer Ordnung beschrieben.

Um die Aberrationen des Auges zu messen, wird ein Laserstrahl auf die Netzhaut fokussiert und das an der Retina diffus reflektierte Licht durch das Linsenarray eines Hartmann-Shack-Sensors geleitet. Aufgrund der inhärenten Bewegung des zu untersuchenden Auges mit einer Frequenz von 80 bis 90 Hz ist die Detektion der Fokuspunkt-Positionen mit einer hohen Geschwindigkeit durchzuführen. Die Intensität des Laserstrahls zur Erzeugung des Fokuspunktes ist dabei durch die medizinisch vorgegebene maximal applizierbare Lichtleistung definiert. Unter Vorgabe einer Reflektivität von ca. 1% entsteht eine Strahlungsleistung des reflektierten Lichtstrahls in einer Größenordnung von wenigen μW .

Die beiden genannten Spezifikationen der erforderlichen Wiederholrate und optischen Sensitivität motivieren die Entwicklung eines ASIC als Fokuspunkt-Positiondetektor. Durch die gemeinsame Anordnung von Photodetektoren und Signalverarbeitung auf demselben integrierten Schaltkreis ist es möglich, die Detektion der Fokuspunkt-Positionen bei einer geringen Lichtleistung mit einer hohen Wiederholrate zu realisieren. In dieser Arbeit wird die Spezifizierung, die Entwicklung und die Messung eines Wellenfrontsensor-ASIC vorgestellt.

In Kapitel 1 wird das System zur Mikromorphometrie der Retina erläutert. Das Modell der aktiven Aberrationskompensation wird skizziert und aus den erforderlichen Eigenschaften des Modells die Spezifikation zur Entwicklung des ASIC abgeleitet. Die mathematische Beschreibung der

optischen Aberrationen durch ein charakteristisches Polynom und ein Verfahren zur Berechnung der Koeffizienten dieses Polynoms geben Aufschluss über die Algorithmik der Mustererkennung, die durch den ASIC realisiert wird.

Kapitel 2 gibt eine Übersicht über die optischen und elektronischen Komponenten eines ASIC. Die spektrale Empfindlichkeit der verfügbaren Photodetektoren wird definiert und das Verhalten der aktiven Elemente in der Verarbeitung von geringen Photoströmen untersucht. Aus diesen Erkenntnissen erfolgt die Entwicklung des Wellenfrontsensor-ASIC in Kapitel 3. Der Fokus des Kapitels liegt in der Beschreibung der implementierten analogen Signalverarbeitung, durch die eine schnelle Detektion von Fokalkpunkt-Positionen bei geringen Lichtleistungen erreicht wird. Weiterhin wird ein Modul des Wellenfrontsensor-ASIC vorgestellt, mit dem eine Datenkompression und eine Übertragung der Positionsdaten an ein externes System erfolgt. Den Abschluss des Kapitels bildet die Vorstellung zweier Design-Studien, aus deren Eigenschaften die Erkenntnisse zur Entwicklung des Wellenfrontsensor-ASIC stammen. Das externe System, durch das die Daten des Wellenfrontsensor-ASIC an ein Datenverarbeitungssystem übertragen wird, ist Gegenstand von Kapitel 4.

Die beiden abschließenden Kapitel stellen die Ergebnisse von Messungen dar. Kapitel 5 fasst die Resultate der Studie eines Augenbewegungs-Verfolgers zusammen, der mit Hilfe des ersten Prototypen realisiert wurde. Ein System wurde aufgebaut, mit dem sich Augenbewegungen mit hoher Auflösung und einer Wiederholrate von 1 kHz messen lassen. In Kapitel 6 werden die Messungen des Wellenfrontsensor-ASIC vorgestellt. Zu Beginn des Kapitels erfolgt eine Beurteilung der Prozessqualität der Herstellung des ASIC. Das Mismatching von Transistorstrukturen wird quantifiziert und die Abweichung der tatsächlich realisierten Transistorparameter von den Simulationsvorgaben dargestellt. Um die Funktionsfähigkeit des ASIC zu bewerten, werden die Ergebnisse von Mismatching- und Rauschanalysen der Signalverarbeitung besprochen. Abschließend erfolgt eine Präsentation der Resultate eines Messsystems, mit dem die optischen Aberrationen von Testlinsen quantifiziert wurden.

Kapitel 1

Auflösungs-Verbesserung der Retina-Mikromorphometrie

Im folgenden Kapitel wird die Motivation zur Entwicklung eines Wellenfrontsensor-ASIC beschrieben. Es erfolgt eine Übersicht über den Aufbau eines industriell gefertigten und im praktischen Einsatz der Ophthalmologie befindlichen Retina-Scan-Tomographen (RST). Die negative Auswirkung der optischen Aberrationen des zu untersuchenden Auges auf die Auflösung des Systems wird besprochen und ein System vorgestellt, dass durch die Verwendung einer aktiven optischen Kompensation eine Verbesserung der Bildauflösung erzielen kann. Dazu erfolgt eine Bestimmung der momentan unter praktischen Bedingungen erreichbaren optischen Auflösung und eine Abschätzung der unter Einsatz einer Aberrations-Kompensation möglichen Auflösungsverbesserung. Ein mathematisches Modell zur Beschreibung einer aberrierten Wellenfront wird vorgestellt, und die Architektur des ASIC wird anhand des Hartmann-Shack-Verfahrens zur Detektion von Wellenfront-Aberrationen besprochen. Zum Abschluss des Kapitels werden die Spezifikationen für die Entwicklung des ASIC festgelegt.

1.1 Anwendung von Retina-Scan-Tomographen

Der Aufbau eines Retina-Scan-Tomographen wird anhand des *Heidelberg Retina Tomographen* (HRT) der Firma *Heidelberg Engineering GmbH* beschrieben. Das Gerät befindet sich seit mehreren Jahren im praktischen Einsatz der Ophthalmologie und erreicht derzeit eine nur durch die Aberrationen des zu untersuchenden Auges begrenzte Auflösung der Mikromorphometrie. Durch das System lässt sich die Oberfläche der Retina dreidimensional visualisieren und es ergeben sich für den Ophthalmologen diagnostische Möglichkeiten, um akute oder potentielle Erkrankungen des Auges zu erkennen. Es ist z.B. der Bereich des Sehnerven-Kopfes (Papille) des Auges von Interesse, da sich dort eine Überhöhung des Augeninnendruckes durch Verformungen der peripapillären topographischen Struktur erkennen lässt. Diese pathologische Veränderung der Papille ermöglicht eine Diagnose der *Glaukom*-Erkrankung, auch als *Grüner Star* bezeichnet, die im fortgeschrittenen Stadium zu einem Absterben der Nervenfasern der Netzhaut und so zu einer Erblindung des Betroffenen führen kann. Die Augenkrankheiten, die unter dem

Begriff “Glaukome” zusammengefasst werden, haben unterschiedliche Erscheinungsformen, im Allgemeinen entstehen sie jedoch durch einen erhöhten Augeninnendruck. Wenn der Abfluss für das Kammerwasser blockiert wird, steigt der Druck im Auge an und durch die verminderte Blutversorgung der Retina werden die Fasern, aus denen der Sehnerv besteht, zerstört. Diese beginnende Zerstörung zeigt die ersten Erscheinungen am Sehnerven-Kopf und kann durch einen RST visualisiert werden.

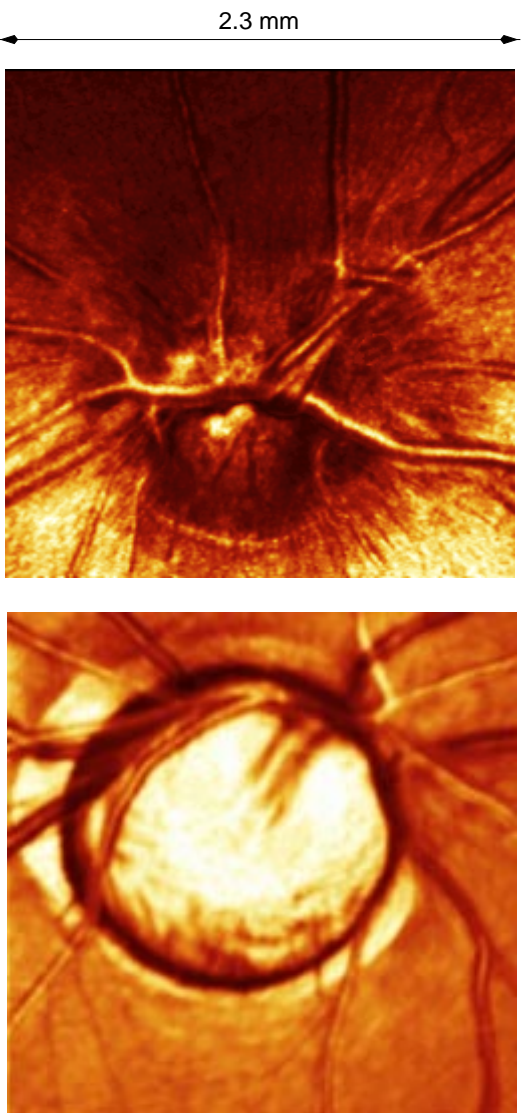


Abbildung 1.1: Visualisierung der Papille der Retina mit einem RST

Abbildung 1.1 zeigt auf der linken Seite eine mit dem HRT gescannte Retina ohne Befund, auf der rechten Seite eine pathologische Veränderung der Papille, verursacht durch erhöhten Augeninnendruck. Bei genauerer Betrachtung der linken Abbildung lässt sich erkennen, dass bei maximaler Auflösung des HRT von ca. $8 - 10 \mu m$ lateral und ca. $220 \mu m$ vertikal makroskopische Strukturen im peripapillären Bereich visualisierbar sind, mikroskopische Elemente, mit einer Strukturgröße von ca. $2 - 3 \mu m$, jedoch nicht mehr darstellbar sind. Im anschließenden Abschnitt wird diese Auflösungsbegrenzung begründet und quantifiziert. Abbildung 1.2 verdeutlicht den Strahlverlauf des abtastenden und des reflektierten Strahls. Der eintretende Strahl ist in dieser Skizze gegenüber dem reflektierten Strahl im Durchmesser zur Übersichtlichkeit geringer dargestellt. In praktischen Anwendungen besitzen beide Strahlen im Allgemeinen den gleichen Durchmesser durch die Pupille begrenzten Durchmesser, wobei Änderungen des Strahldurchmessers des einfallenden Strahls zur verbesserten Detektion der Aberrationen Gegenstand der Forschung sind [Artal 94].

1.2 Die aberrationsbegrenzte Auflösung eines RST

Die in Abbildung 1.1 dargestellte Information wird mit dem HRT durch ein dreidimensionales Abtasten der Retina erzeugt, siehe Abbildung 1.3. Ein Laserstrahl wird durch einen Abtaster, bestehend aus zwei Galvanometer-Scannern, lateral abgelenkt und über ein optisches System zur Vorkompensation von defokalen Fehlsichtigkeiten des Patienten in das zu untersuchende Auge geleitet. Das an der Retina reflektierte Licht folgt dem Eingangs-Strahlengang, wird durch die Galvanometer-Scanner wieder in den ursprünglichen Weg geleitet und über einen Strahlteiler ausgekoppelt. Eine Avalanche-Photodiode empfängt das an der Retina reflektierte Licht und

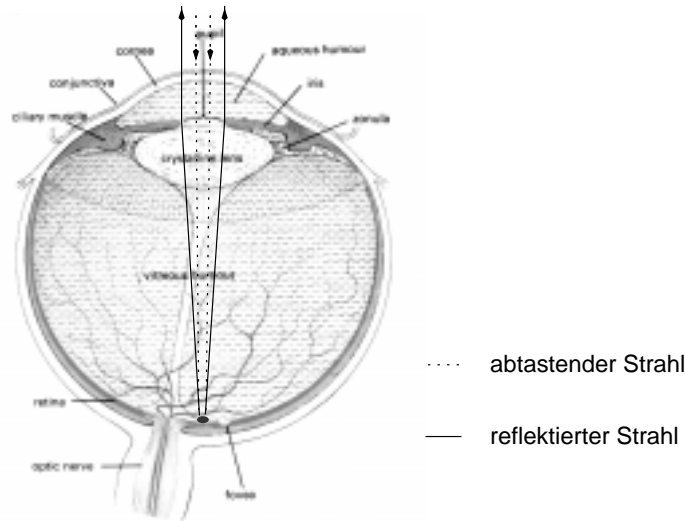


Abbildung 1.2: Strahlverlauf im abgetasteten Auge

erzeugt der Intensität entsprechend einen Photostrom. Zur Realisierung einer dreidimensionalen Abtastung ist das System zusätzlich als konfokales Mikroskop aufgebaut. Zwischen dem Strahlteiler und der Avalanche-Photodiode wird der Strahl auf eine Lochblende fokussiert. Nur das Licht, das an der Oberfläche der Retina in der zur Lochblende konjugiert liegenden Ebene reflektiert wird, ist von der Avalanche-Photodiode detektierbar. Durch Verschiebung des gesamten Abtast-Systems auf die letzte Objektiv-Linse vor dem Auge des Patienten zu, entlang der optischen Achse, kann die Ebene, in der Reflexion detektierbar ist, eingestellt werden. Aus den in Schichten aufgenommenen Daten wird die Topographie der Retina dreidimensional rekonstruiert. Die Auswirkungen der optischen Aberrationen des Auges auf die Auflösung des HRT

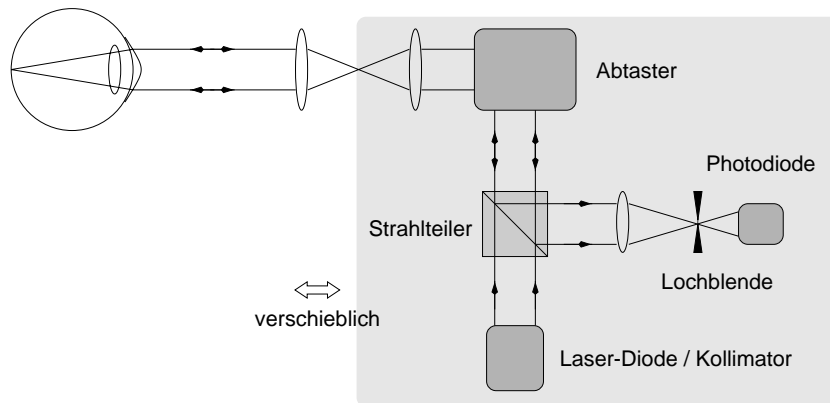


Abbildung 1.3: Schematischer Aufbau des HRT

wird für die laterale und die axiale Richtung unterschieden. In lateraler Richtung bestimmt theoretisch die Ausdehnung der Pupille die Aufweitung des Fokalkpunktes auf der Retina. Nach den Gesetzen der Fourier-Optik [Bergmann 92] ergibt sich mit

$$d = 1.22 \cdot \frac{\lambda f}{D} \quad (1.1)$$

für die Fokallänge des Auges $f = 17\text{ mm}$, dem Durchmesser der Pupille von $D = 6\text{ mm}$, und der verwendeten Lichtwellenlänge $\lambda = 780\text{ nm}$ ein Durchmesser des Fokalkpunktes von $d = 2.7\text{ }\mu\text{m}$. Praktische Messungen zeigen jedoch, dass aufgrund der bei dieser Pupillengröße verstärkt auftretenden Aberrationen ein vergrößerter Durchmesser entsteht. Nach [Campbell 66] beträgt der zur Vermeidung von optischen Aberrationen geeignete Pupillen-Durchmesser maximal 2.4 mm . Bei dieser Größe ergibt sich ein Durchmesser des Fokalkpunktes von theoretisch $d = 5.5\text{ }\mu\text{m}$. In dem System des HRT wird derzeit ein Spot-Durchmesser von ca. $8 - 10\text{ }\mu\text{m}$ erzielt, was zum Teil auch auf die durch die Blend-Wirkung des Patienten verursachte Verkleinerung der Pupille zurückzuführen ist.

Die axiale Auflösung bestimmt sich durch eine Beurteilung der konfokalen Anordnung des Systems. Dabei wird die Halbwertsbreite der Intensitätsverteilung des Fokalkpunktes in Richtung der optischen Achse an der Lochblende bestimmt. Diese ist eine Funktion des Abstandes des reflektierenden Objektes (Retina-Schicht) von der Fokal-Ebene der Objektiv-Linse (Hornhaut / Linse). Die Lochblende trägt dabei die Funktion eines geometrischen Raumfilters, der eine durch den Radius der Blende definierte Integration der Intensitäts-Verteilung innerhalb der Grenzen der Lochblende ermöglicht. Die Tiefenauflösung ist definiert als das Maß der Verschiebung der Fokalebene, durch welches einen Abfall der maximalen Intensität um 50% hinter der Lochblende verursacht. Die Berechnung der Tiefenauflösung ist ausführlich in [Born 65] dargestellt, zur Übersichtlichkeit sei an dieser Stelle nur das Ergebnis der Berechnung angeführt. Die theoretisch erreichbare Tiefenauflösung des HRT beträgt bei einer geweiteten Pupille des Patienten von $D = 6\text{ mm}$ ca. $40\text{ }\mu\text{m}$, Beugungsbegrenzung vorausgesetzt. Es wurde in praktischen Messungen jedoch bei diesem Pupillen-Durchmesser eine Tiefenschärfe von $487 \pm 75\text{ }\mu\text{m}$ festgestellt [Dreher 89]. Bei einem Pupillen-Durchmesser von $D = 2.5\text{ mm}$ konnte eine Tiefenauflösung von $332 \pm 62\text{ }\mu\text{m}$ festgestellt werden, im Vergleich zu der theoretisch erreichbaren Auflösung von $206\text{ }\mu\text{m}$. Die im Allgemeinen erreichbare Tiefenauflösung des HRT beträgt ca. $220\text{ }\mu\text{m}$.

1.3 Auflösungsverbesserung eines RST durch aktive Optik

Die erreichbare optische Auflösung eines RST ist bei geweiteten Pupillen mit einem Durchmesser von $D > 2.5\text{ mm}$ aberrationsbegrenzt. Die sich bei Vergrößerung der Pupille verringern den Beugungserscheinungen werden durch das Auftreten optischer Aberrationen von Linse und Hornhaut überlagert. Es ist daher sinnvoll, die Aberrationen des Linsen/Hornhaut-Systems zu kompensieren, um bei erhöhter Pupillengröße eine Beugungsbegrenzung zu erzielen. Die optische Auflösung eines RST würde sich dadurch verbessern. Der Aufbau eines solchen aberrationskompensierenden Systems ist das Ziel der Entwicklung eines Wellenfrontsensor-ASIC und eines deformierbaren Spiegels¹. Das ASIC-System detektiert die Aberrationen des zu untersuchenden Auges und erzeugt die Koeffizienten eines Polynoms höherer Ordnung zur Beschreibung der Wellenfront. Die Daten werden an ein System übergeben, das anhand der Koeffizienten einen deformierbaren Spiegel steuert. Dieser Spiegel bringt eine ortsabhängige Phaseninformation in den das Auge abtastenden Laser-Strahl ein, die durch die Aberrationen des Auges kompensiert wird. Der abtastende Fokalkpunkt erhält dadurch eine beugungsbegrenzte Gestalt, die Auflösungsverbesserungen, die in Kapitel 1.2 erörtert wurden, können ermöglicht werden. In Abbildung 1.4 ist das System des HRT, erweitert um die Einheit der Aberrationskompensation, dargestellt.

Der Prozess der Aberrationskompensation erfordert ein iteratives Steuern und Messen der erzeugten Kompensations-Qualität. Da zu Beginn der Kompensation das Maß der Aberrations-

¹Der Einsatz des deformierbaren Spiegels wird innerhalb des Projektes ebenfalls an der Universität Heidelberg erforscht

Zum anderen verursachen Fehler in der mechanischen Form der optischen Elemente Aberrationen, die in ihren Auswirkungen besonders in der Ophthalmologie von Interesse sind. Abweichungen von der idealen Krümmungsform von Hornhaut und Linse verursachen Aberrationen, wobei im menschlichen Auge besonders die defokalen oder astigmatischen Aberrationen Fehlsichtigkeiten verursachen. Bei Vergrößerung des Pupillen-Durchmessers treten zusätzlich asymmetrische (Koma) oder sphärische Aberrationen auf. In der Dämmerung entstehen durch die zur Adaption an die geringere Licht-Intensität geweitete Pupille stärkere Aberrationen von Linse und Hornhaut, so dass eine Verschlechterung des Sehvermögens eintritt.

Eine physikalische Begrenzung in der Verbesserung der Qualität von optischen Elementen ist durch Beugung gegeben. Die Beugungsbegrenzung ist eine Konsequenz der nicht unendlichen Ausdehnung der begrenzenden Aperturen und entsteht in allen optischen Systemen zwangsläufig. Es ist wünschenswert, die Abbildungseigenschaften des Systems nur auf die durch Beugung entstehenden Abbildungsfehler zu reduzieren. Bei der Untersuchung des menschlichen Auges wird die Pupille oft künstlich geweitet (dilatiert), um eine möglichst hohe Intensität der Reflexion im rückwärtigen Strahlengang zu ermöglichen und die Beugungserscheinungen zur Verbesserung der Auflösung zu reduzieren. Durch diese Weitung treten jedoch die optischen Aberrationen höherer Ordnungen (sphärische Aberrationen, Koma) auf. Man spricht von einer *aberrationsbegrenzten* Auflösung des Visualisierungs-Verfahrens. Solange die Pupille einen Durchmesser von 2.4 mm nicht überschreitet, wird im Allgemeinen Beugungsbegrenzung angenommen [Campbell 66]. Ziel des Systems der aktiven Aberrations-Kompensation ist die Realisierung der ausschließlich *beugungsbegrenzten* Auflösung bei einem Pupillen-Durchmessers von mehr als 2.4 mm .

Nachfolgend wird ein mathematisches Modell zur Beschreibung von aberrierten Wellenfronten vorgestellt und typische Formen von optischen Aberrationen und ihre Auswirkungen auf die Gestalt der Wellenfront des sich ausbreitenden Lichtstrahles beschrieben.

1.4.1 Die Zernike-Polynome

Eine Form zur Beschreibung der Ausbreitung elektrischer Felder ist die Wellengleichung. Die Feldstärke des elektrischen Feldes im verlustfreien Medium folgt der Gleichung

$$\vec{E} = \vec{E}_0 \cos(2\pi\nu t - \vec{k}\vec{r} + \phi), \quad (1.2)$$

mit der elektrischen Feldstärke E_0 , dem Ortsvektor $\vec{r} = (x, y, z)$, dem Ausbreitungsvektor \vec{k} in Richtung \vec{r} , der Phase ϕ , der Frequenz ν und der Zeit t . Dabei sind die Flächen konstanter Phase gegeben durch

$$\vec{k}\vec{r} - 2\pi\nu t = \text{konstant}. \quad (1.3)$$

Von besonderem Interesse für die Definitionen von Aberrationen ist die Phase ϕ . In Gleichung 1.2 wird eine ortsunabhängige Phase senkrecht zu dem Ausbreitungsvektor k angenommen. Im Folgenden wird diese Phase jedoch nicht mehr als konstant, sondern als abhängig von der lateralen Entfernung senkrecht zu k angenommen:

$$\phi = \Phi(x, y) \quad (1.4)$$

Die Phase sei nun $\Phi(x, y)$; analog zu der Definition der Phase in Bezug auf kartesische Koordinaten können auch Polarkoordinaten verwendet werden: $\Phi(\rho, \phi)$. Die Einheit der Phase ist $[rad]$. Es wird nun eine auf die verwendete Wellenlänge λ bezogene Größe der Wellenfront $W(x, y)$ definiert:

$$W(x, y) \equiv \frac{\lambda}{2\pi} \Phi(x, y) = [m]. \quad (1.5)$$

Die Wellenfront $W(x, y)$ ist zu verstehen als eine Fläche von Orten gleicher Phase und wird mathematisch beschrieben durch die Differenz der optischen Pfadlänge (*optical path difference (OPD)*) zwischen der realen Wellenfront und der idealen flächenartigen Wellenfront senkrecht zum k -Vektor, an jeder Stelle (x, y) , unter Verwendung der Einheit $[m]$.

Da real auftretende Wellenfronten stetig differenzierbar sind, ist es möglich, eine zweidimensionale Funktion höherer Ordnung zu entwickeln, die die oben besprochene Differenz $W(x, y)$ oder $W(\rho, \phi)$ mit wählbarer Genauigkeit annähern kann. In der Literatur werden verschiedene Funktionen zitiert, die zumeist eine Summe mehrgradiger Polynome mit jeweiligen Koeffizienten repräsentieren. Eine in der Optik und der Ophthalmologie häufig verwendete Darstellung wurde von F. Zernike 1934 [Zernike 34] zur Beschreibung der Qualität von Hohlspiegeln entwickelt. Die Zernike-Polynome besitzen Orthogonalität, den typischen Formen von Aberrationen wie Defokus, Astigmatismus, Koma und sphärische Aberrationen lassen sich dedizierte Koeffizienten zuordnen (siehe Tabelle 1.1). Zernike beschreibt eine Wellenfront als Summe aus den Produkten dieser Polynome $Z_i(x, y)$ mit den Koeffizienten C_i :

$$W(x, y) = \sum_{i=1}^n C_i \cdot Z_i(x, y) \quad (1.6)$$

C_i	Z_i (polar)	Z_i (kartesisch)	Bedeutung
C_1	$\rho \sin \phi$	x	Verkipfung in X-Richtung
C_2	$\rho \cos \phi$	y	Verkipfung in Y-Richtung
C_3	$\rho^2 \sin 2\phi$	$2xy$	Astigmatismus in $\pm 45^\circ$
C_4	$2\rho^2 - 1$	$-1 + 2y^2 + 2x^2$	Defokus
C_5	$\rho^2 \cos 2\phi$	$y^2 - x^2$	Astigmatismus in 0° or 90°
C_6	$\rho^3 \sin 3\phi$	$3xy^2 - x^3$	Koma in X-Richtung Koma in Y-Richtung
C_7	$3(\rho^3 - 2\rho) \sin \phi$	$-2x + 3xy^2 + 3x^3$	
C_8	$3(\rho^3 - 2\rho) \cos \phi$	$-2y + 3y^3 + 3x^2y$	
C_9	$\rho^3 \cos 3\phi$	$y^3 - 3x^2y$	Sphärische Aberrationen
C_{10}	$\rho^4 \sin 4\phi$	$4y^3x - 4x^3y$	
C_{11}	$(4\rho^4 - 2\rho^2) \sin 2\phi$	$-6xy + 8y^3x + 8x^3y$	
C_{12}	$6\rho^4 - 6\rho^2 + 1$	$1 - 6y^2 - 6x^2 + 6y^4 + 12x^2y^2 + 6x^4$	
C_{13}	$(4\rho^4 - 2\rho^2) \cos 2\phi$	$-3y^2 + 3x^2 + 4y^4 - 4x^4$	
C_{14}	$\rho^4 \cos 4\phi$	$y^4 - 6x^2y^2 + x^4$	

Tabelle 1.1: Die Zernike-Koeffizienten und -Polynome

In Tabelle 1.1 sind die Koeffizienten bis zur 4. Ordnung dargestellt, wobei auch höhere Ordnungen beschrieben werden können [Malacara 92]. In Abbildung 1.5 werden typische Beispiele von Aberrationen gezeigt. Die Berechnung des jeweiligen dreidimensionalen Körpers und des dazugehörigen Höhenlinienmodells erfolgt unter Setzen eines einzelnen Koeffizienten zu $C_i = 1 \mu m$ und aller anderen Koeffizienten zu $C_n = 0$, $n \in \{1, 2, \dots, 14\} \setminus i$. Die Höhenlinienmodelle aus Abbildung 1.5, b) stellen die Iso-Höhenlinien der Wellenfronten aus a) für $x, y \in (-0.71, 0.71)$ dar. Die Abstände der benachbarten Iso-Höhenlinien sind neben der Grafik als d angegeben. Die Ortsvariablen x und y werden im Allgemeinen auf den die vorhandenen Aperturen umschreibenden Einheitskreis normiert und sind daher einheitenlos.

Da die Ortsvariablen x und y einheitslos sind, ist es notwendig, bei der Angabe von Zernike-Polynomen auch den Durchmesser oder den Radius der Apertur anzugeben, auf den die Koeffizienten bezogen werden. In manchen Fällen werden die Koeffizienten in $[\frac{1}{m}]$ angegeben, um eine Umrechnung von Koeffizienten in Dioptrien und umgekehrt zu erleichtern. Es ist ebenfalls möglich, die Zernike-Koeffizienten in Einheiten von Lichtwellenlängen λ anzugeben, wobei dann jedoch die quantitative Angabe von λ zum weiteren Verständniss erfolgen muss. Als

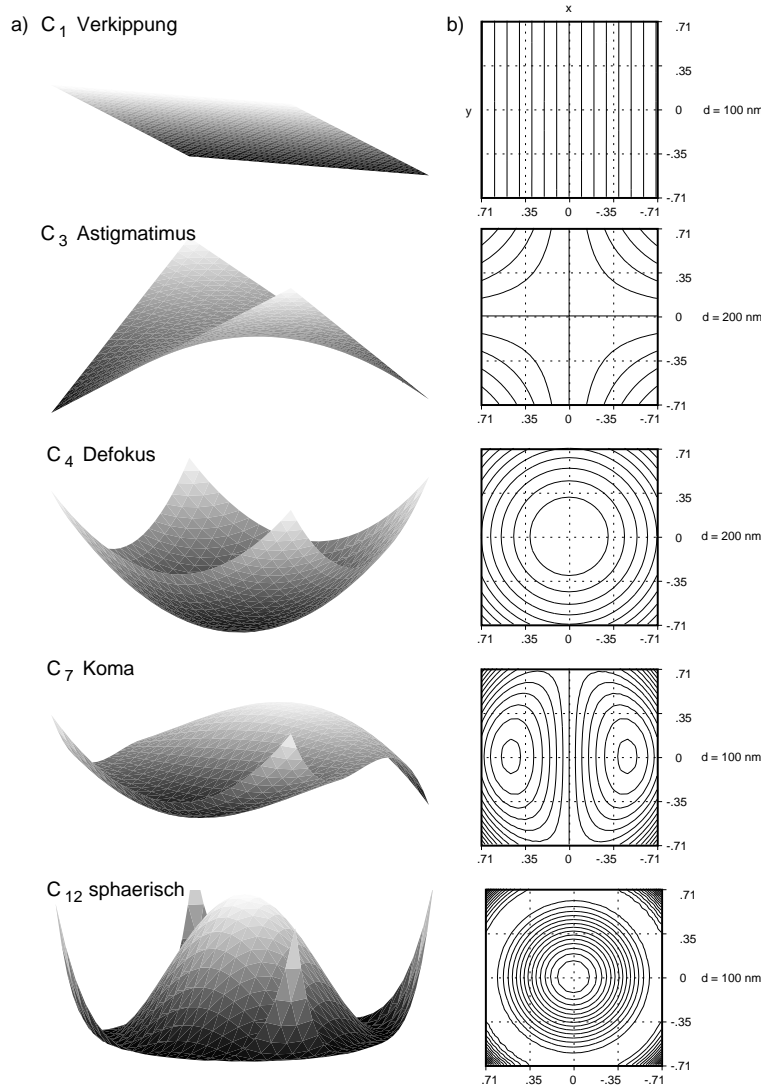


Abbildung 1.5: Beispiele von Wellenfront-Aberrationen

C_i	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}
$[\mu m]$	0.89	2.78	0.43	0.20	-0.32	0.23	-0.29	-0.09	0.10	0.15	-0.17	0.03

Tabelle 1.2: Zernike Koeffizienten aus [Liang 94]

Beispiel einer durch Aberrationen eines menschlichen Auges entstehenden Wellenfront sei an dieser Stelle ein Bezug auf [Liang 94] gegeben. Die dort gemessene Wellenfront ist mit den in Tabelle 1.2 aufgelisteten Zernike-Koeffizienten beschrieben worden. Abbildung 1.6 stellt diese

Zernike-Koeffizienten als dreidimensionalen Körper und Höhenlinienmodell dar. Es sei angemerkt, dass die Koeffizienten in $[\mu m]$ notiert wurden, und sich auf eine Apertur der Pupille des gemessenen Auges von 6 mm beziehen. Die Kurzsichtigkeit des Auges wurde mit -1.0 bis -1.5 Dioptrien angegeben. Im oberen Teil der Grafik wurde die komplette Wellenfront dargestellt, im unteren Teil die Wellenfront nach Setzen von $C_4 = 0$ zur Unterdrückung der defokalen Aberration. Alle anderen Aberrationen bleiben als eine Auswirkung der Orthogonalität der Zernike-Polynome erhalten. Dadurch lassen sich die höherwertigen und astigmatischen Aberrationen, die im Allgemeinen einen geringeren dynamischen Bereich als die defokalen Aberrationen besitzen, hervorheben.

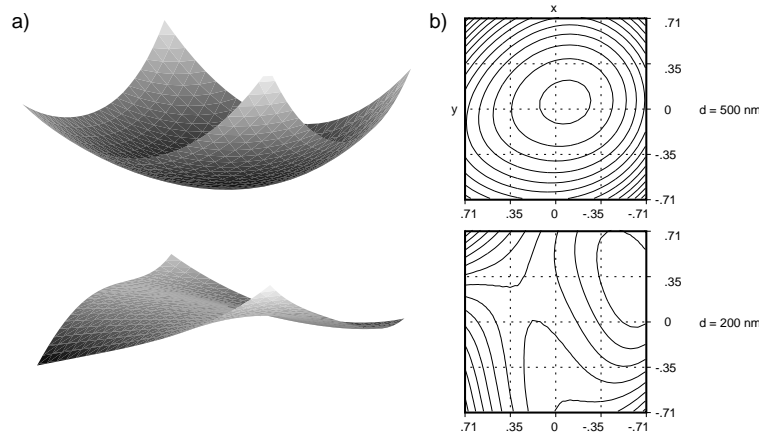


Abbildung 1.6: Aberrationen eines menschlichen Auges mit und ohne defokalem Anteil [Liang 94]

1.5 Bestimmung einer Wellenfront

Für das in Kapitel 1.4 besprochene mathematische Modell einer Wellenfront werden im Folgenden Messverfahren und Rekonstruktionsalgorithmen erläutert. Die Rekonstruktion einer Wellenfront-Information unterliegt zwei aufeinander folgenden Schritten, der Quantifizierung der Phasendifferenzen zwischen einer Referenzwellenfront und der zu messenden Wellenfront, und der Berechnung der Koeffizienten zur Modellierung der erhaltenen Informationen der Phasendifferenz durch Polynome. Im Folgenden werden verschiedene Prinzipien zur Messung dieser Phasendifferenzen vorgestellt und ein Verfahren als Ausgangsprinzip zur Definitionen der ASIC-Spezifikationen in Kapitel 1.6 besonders hervorgehoben. Anschließend wird ein mathematisches Schema zur Entwicklung der Koeffizienten der Zernike-Polynome aus den Phasendifferenzinformationen beschrieben.

Die genaueste Messung von Phasendifferenzen erfolgt mit interferometrischen Systemen. Dazu werden die durch Überlagerungen von zu messender Wellenfront und Referenzwellenfront entstehenden Intensitätsmaxima und -minima erfasst und ausgewertet. Durch Verschieben der Referenzwellenfront oder der Messwellenfront in geringer Schrittweite lassen sich Feinphasendiagramme erzeugen, die Genauigkeiten der Phasendifferenz-Bestimmung von einigen Bruchteilen der Lichtwellenlänge der verwendeten Beleuchtungsquelle erlauben. Zu dieser Feinphasen-Interferometrie werden jedoch über einen Zeitraum wiederholt Interferogramme erstellt und ausgewertet. Innerhalb der gesamten Messphase dürfen weder die Referenz- noch die Messwellenfront Änderungen unterliegen. Eine praktische Anwendung in der Ophthalmologie ist nicht

möglich.

Zur Bestimmung der Phasendifferenzen aufgrund von Einzelmessungen werden daher bevorzugt zwei Verfahren angewandt, die Rückprojektion der Wellenfront durch Auswertung der *Point Spread Function (PSF)*, also der lateralen Intensitätsverteilung eines Fokalkpunktes, und die Bestimmung der lokalen Wellenfront-Verkipfung in Subaperturen durch Detektion von Fokalkpunkt-Verschiebungen. In der Auswertung der PSF einer fokussierten Wellenfront wird über eine Fouriertransformation die Phasendifferenzinformation vor der Fokussierung zurückgewonnen [Artal 88]. Dieses Verfahren erlaubt eine sehr genaue Bestimmung der Wellenfrontaberrationen, bedarf aufgrund der komplexen Berechnungs-Algorithmen jedoch einer hohen Rechenleistung. Ein ASIC ist auf dieses Verfahren der Wellenfrontrekonstruktion nicht adaptierbar, da die optische Auflösung, mit der die fokale Abbildung erfasst werden muss, nicht realisierbar ist. Weiterhin ist fraglich, ob die Datenverarbeitungs-Geschwindigkeit zur Einhaltung der Zeit-Spezifikationen des ASIC erreicht werden kann.

Das Hartmann-Shack-Verfahren der Wellenfrontdetektion, das Grundlage der Entwicklung des Wellenfrontsensor-ASIC ist, lässt sich als Detektionsschema in dem System zur Auflösungsverbesserung der Mikromorphometrie einsetzen. Es bestimmt aus der Verschiebung eines Musters von Fokalkpunkten die lokale Verkipfung von Wellenfront-Subaperturen. Das Verfahren ist eine parallelisierte Anwendung des Hartmann-Verfahrens [Geary 95] und wird seit der Möglichkeit zur Entwicklung von hochqualitativen Mikrolinsen oft verwendet. Aktive optische Systeme zur Kompensation der atmosphärischen Störungen in astronomischen Teleskopen verwenden bevorzugt dieses Detektionsverfahren [Merkle 88]. Es folgt eine Beschreibung des Verfahrens.

1.5.1 Das Hartmann-Shack-Detektionsverfahren

Eine Anordnung von Mikrolinsen unterteilt die Apertur des zu messenden Lichtstrahls in Subaperturen. Jede Linse erzeugt einen Fokalkpunkt im Abstand der Fokallänge der Linse. Bei einer lokalen Verkipfung der Wellenfront innerhalb der Linsenapertur erfolgt eine Verschiebung des Fokalkpunktes aus der Position x_0 nach x^* um den Betrag Δx . In erster Näherung ist diese Verschiebung genau proportional der Verkipfung der Wellenfront innerhalb der Linse. Abbildung 1.7 a) zeigt diese Verschiebung. Ein Hartmann-Shack-Wellenfrontdetektor erfasst diese Verschiebung und schliesst dadurch auf das Maß der Verkipfung der Wellenfront innerhalb der Linse. Nach entsprechender Detektion einer zweidimensionalen Matrix von Fokalkpunkt-Verschiebungen kann die komplette Wellenfront rekonstruiert werden. Mögliche Fehler bei der Rekonstruktion der Wellenfrontverkipfung innerhalb einer Linse entstehen dabei durch die Unsicherheiten in der Bestimmung der Fokalkpunkt-Verschiebung. Die Bestimmung der Verschiebung wird im Folgenden genauer betrachtet.

Nach den Vereinfachungen der Paraxial-Optik² kann der folgende Zusammenhang definiert werden:

$$\frac{dW}{dx} = \frac{\Delta x}{f} = \alpha \quad (1.7)$$

Dabei ist dW/dx die Verkipfung der Wellenfront innerhalb der Linse, Δx die Verschiebung des Fokalkpunktes und f die Brennweite der Linse. Praktische Größen verdeutlichen dabei die Legitimation der Annahme der Gauß-Optik: $f = 50 \text{ mm}$, $dW/dx = \alpha = 200 \text{ nm}/400 \mu\text{m} = 5 \cdot 10^{-4}$, $\Delta x = 25 \mu\text{m}$. Zur genaueren Betrachtung sei angemerkt, dass die Intensitätsverteilung des Fokalkpunktes einer Gauß'schen Verteilung gleicht, siehe Abbildung 1.7 b) (das Bild entspricht der praktischen Intensitätsverteilung, erzeugt mit einer im System verwendeten Mikrolinse, aufgenommen durch eine CCD-Kamera). Theoretisch existiert genau ein Maximum der Intensität

² $\sin(x) = x$

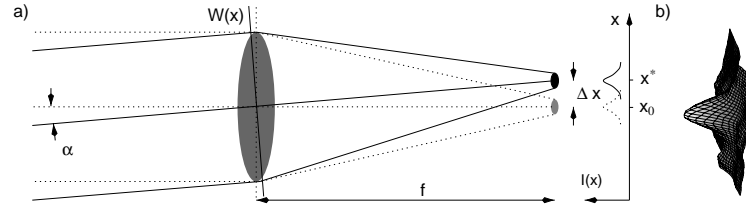


Abbildung 1.7: Verschiebung des Fokalfunktes bei verkippter Wellenfront

an einer Stelle x^* . Es ist für einen Detektor praktisch jedoch nicht möglich, genau diese Stelle x^* zu erfassen, da das Detektionsverfahren eine Diskretisierung der Abszisse vornimmt und nur das Intensitäts-Mittel eines Abschnittes Δx_d detektiert. Um die Güte der Detektion zu qualifizieren, muss daher eine Abschätzung der Detektions-Genauigkeit in Bezug auf diese Breite Δx_d erfolgen (siehe Kapitel 1.6.1). Die Breite der Intensitätsverteilung des Fokalfunktes ist für die Genauigkeit der Detektion der Fokalfunkt-Verschiebung eklatant. Nach den Gesetzen der Fourier-Optik beträgt die Breite des Fokalfunktes (Beugungs-Scheibchen)

$$D_{FP} = 1.22 \cdot \frac{\lambda f}{D_L}. \quad (1.8)$$

Mit dem Linsendurchmesser $D_L = 400 \mu m$, der Wellenlänge $\lambda = 780 nm$ und der Fokallänge $53 mm$ ergibt sich ein Durchmesser des ersten Intensitätsminimum-Ringes des Fokalfunktes von $126 \mu m$. Ein Detektor muss zur Bestimmung des Maxima mit einer Auflösung von ca. $15 - 20 \mu m$ eine Diskriminierungsgenauigkeit der Intensitätsunterschiede von wenigen Prozent aufweisen. Diese Anforderung ist bei entsprechender Verringerung des Durchmessers des Fokalfunktes einfacher zu realisieren. Ebenfalls muss beachtet werden, dass die Form der Wellenfront innerhalb der Linse im Allgemeinen nicht einer konstanten Verkipfung gleicht, sondern vielmehr eine Unregelmässigkeit besitzt, die eine weitere Aufweitung des Fokalfunktes und damit eine Unsicherheit in der Bestimmung der Verschiebung Δx verursacht. Ebenso werden die Auswirkungen der optischen Aberrationen der Linse selbst Einfluss auf die Form des Fokalfunktes haben. Diese Problematiken werden während der Aufstellung der Spezifikationen für die Entwicklung des ASIC in Kapitel 1.6.2 genauer betrachtet.

Der Hartmann-Shack-Detektor verwendet zur Rekonstruktion der Wellenfront mehrere Linsen, die äquidistant in einer zweidimensionalen lateralen Anordnung im Detektor-Strahlengang angebracht werden. Dieses *Linsenarray* erzeugt eine Matrix von Fokalfunkten. Die Verschiebung jedes Fokalfunktes ist ein Maß für die Verkipfung der Wellenfront innerhalb der zugehörigen Linse, und aus den Messungen der einzelnen Verschiebungen lässt sich anhand eines Näherungs-Algorithmus' die gesamte Wellenfront annähern. Abbildung 1.8 a) stellt eine 3×3 -Matrix eines Linsenarray dar. Die Fokalfunkte treffen in der Fokal-Ebene auf eine Fläche, in der später der ASIC angeordnet wird. Zur Übersichtlichkeit der Grafik werden nur 3×3 Punkte dargestellt, das System des Wellenfrontsensor-ASIC verwendet eine Matrix von 16×16 Linsen. In Abbildung 1.8 b) ist das Muster der Fokalfunkte *vor* (graue Punkte) und *nach* (schwarze Punkte) der Verschiebung verdeutlicht. Jede Verschiebung $\Delta x_{n,m}$ und $\Delta y_{n,m}$, $n, m \in \{1, 2, 3\}$ wird erfasst und ausgewertet. Die Auswertung erfolgt durch einen *Least-Square-Fit*, der im Folgenden beschrieben wird.

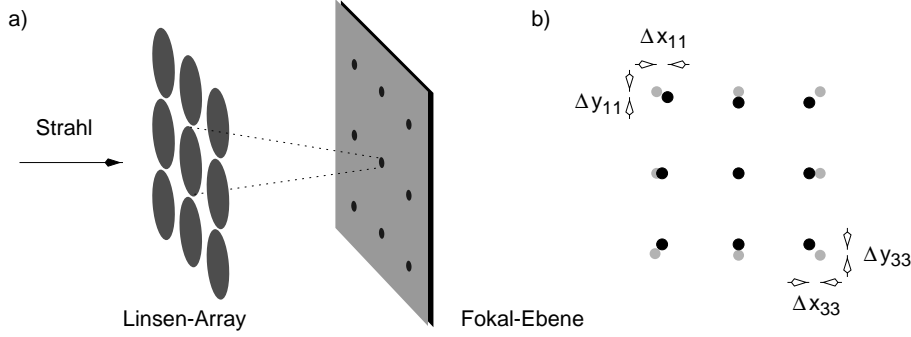


Abbildung 1.8: Das Linsenarray erzeugt eine Matrix von Fokalfunkten

1.5.2 Die Rekonstruktion der Wellenfront

Aus den Informationen der Verschiebungen $\Delta x_{n,m}$ und $\Delta y_{n,m}$, $n \in \{1, 2, \dots, n_{max}\}$, $m \in \{1, 2, \dots, m_{max}\}$ lassen sich die Verkippungen der Wellenfront innerhalb jeder Linse bestimmen:

$$\begin{aligned} \left. \frac{dW}{dx} \right|_{n,m} &= \left. \frac{\partial W(x, y)}{\partial x} \right|_{n,m} = \frac{\Delta x_{n,m}}{f} \\ \left. \frac{dW}{dy} \right|_{n,m} &= \left. \frac{\partial W(x, y)}{\partial y} \right|_{n,m} = \frac{\Delta y_{n,m}}{f} \end{aligned} \quad (1.9)$$

Ausgehend von der mathematischen Beschreibung einer Wellenfront durch ein zweidimensionales Polynom (Zernike-Polynome, siehe Gleichung 1.6) müssen die Koeffizienten der Polynome derart bestimmt werden, dass die partiellen Ableitungen (die Verkippung) $\partial W/\partial x$ und $\partial W/\partial y$ der Summen-Funktion an der Stelle x_n, y_m möglichst genau den Werten $\Delta x_{n,m}/f$ und $\Delta y_{n,m}/f$ entsprechen. Dieses Problem der Anpassung kann durch den Least-Square-Fit-Algorithmus gelöst werden. Zu jedem Fokalfunkt wird die Differenz zwischen den Messwerten und den Werten der Summen-Funktion quadriert. Die Koeffizienten der Polynome werden nun derart berechnet, dass die Summe der Fehlerquadrate von allen Punkten einen minimalen Betrag darstellt. Unter Berücksichtigung der partiellen Ableitungen der Zernike-Polynome nach x, y ergibt sich daher die folgende Beziehung (die Matrizen-Indizes n, m werden durch einen Lauf-Index i ersetzt, weiterhin sei $N = n \times m$):

$$S = \sum_{i=1}^N \left(\left. \frac{\partial}{\partial x} W(x, y) \right|_{x=x_i, y=y_i} - \frac{\Delta x_i}{f} \right)^2 + \sum_{i=1}^N \left(\left. \frac{\partial}{\partial y} W(x, y) \right|_{x=x_i, y=y_i} - \frac{\Delta y_i}{f} \right)^2 \quad (1.10)$$

S wird nun in Bezug auf die Koeffizienten der Zernikepolynome C_i , $i \in \{1, 2, \dots, 14\}$ minimiert, so dass sich für einen Satz von Koeffizienten C_i eine minimale Abweichung der mathematischen Beschreibung der Wellenfront an den Stützstellen x_i und y_i zu den Messwerten ergibt. Die Herleitung des Least-Square-Fit ist in Anhang A beschrieben.

1.6 Anforderungen an den Wellenfrontsensor-ASIC

Die in den vorangehenden Kapiteln beschriebenen System-Eigenschaften, das Detektions-Verfahren und der Berechnungs-Algorithmus definieren in Verbindung mit den Eigenschaften der Anwendung die Spezifikationen zum Entwurf des Wellenfrontsensor-ASIC. Es lassen sich die Anforderungen an die Auflösung und den dynamischen Bereich der Detektion der Fokalfunkte-Verschiebungen, sowie an das zeitliche Verhalten der Detektion spezifizieren.

1.6.1 Auflösung und dynamischer Bereich der Detektion

Der Wellenfrontsensor-ASIC bestimmt die optischen Aberrationen des menschlichen Auges. Es ist notwendig, die möglichen Aberrationsformen abzuschätzen, um den dynamischen Bereich und die erforderliche Genauigkeit der Messung quantifizieren zu können. Die typischen Aberrationen des Auges wurden in Kapitel 1.4 mit Tabelle 1.2 vorgestellt. Hauptsächliche Aberrationstypen sind Defokus und Astigmatismus. Der Astigmatismus ist im Allgemeinen geringer als der Defokus. Astigmatismen von 0 bis ca. 1 *dpt* sind typische Werte; defokale Aberrationen liegen in einem Bereich von 0 bis ca. 12 *dpt*. Das System des HRT ist in der Lage, defokale Aberrationen durch Verschieben einer Objektiv-Linse entlang der optischen Achse in Schritten von 0.25 *dpt* zu kompensieren. Kleinere defokale Anteile der Aberrationen von ca. 0.25 *dpt* bleiben jedoch bestehen. Der Wellenfrontsensor-ASIC muss daher Aberrations-Größen von ca. 0 – 1 *dpt* detektieren können. Nach $C_4^* = \frac{1}{4} D_A^3$ und

$$\frac{\Delta x}{f} = \frac{dW}{dx} = 4 \cdot C_4^* \cdot x \quad (1.11)$$

ergibt sich für eine Aberration von $D_A = \pm 1.0$ *dpt* Hyper- oder Myopie bei einer Apertur der Pupille von $D_P = 6.4$ mm, einer Fokallänge des Linsenarray von $f = 53$ mm und einer Apertur der Linsen von 400 μm in den äußersten Subaperturen der Matrix eine Fokalfunkt-Verschiebung von $\Delta x = \pm 169$ μm . Eine Auflösung der Verschiebungsdetektion von $\Delta x_d = 16.9$ μm würde die Forderung nach einer absoluten Genauigkeit von $\Delta D_A = 0.1$ *dpt* erfüllen, da die Umrechnung der detektierten Verschiebungen in Zernike-Koeffizienten durch den Least-Square-Fit einer proportionalen Fehler-Fortpflanzung unterliegt.

D_A [dpt]	Δx_i [μm], $i \in \{1, 2, \dots, 16\}$															
1.0	-170	-149	-127	-106	-85	-64	-42	-21	21	42	64	85	115	127	149	170
0.1	-17	-15	-13	-11	-8	-6	-4	-2	2	4	6	8	11	13	15	17

Tabelle 1.3: Verschiebungen der Fokalfunkte bei 1.0 *dpt* und 0.1 *dpt* Defokus

Tabelle 1.3 stellt die Verschiebungen der Fokalfunkte des Linsenarray bei Auftreten einer Wellenfront mit einer defokalen Aberration von $D_A = 1$ *dpt* und $D_A = 0.1$ *dpt* dar. Die Fokallänge des Linsenarray beträgt $f = 53$ mm, die Apertur der Pupille $D_P = 6.4$ mm. Es zeigt sich, dass eine Bestimmung des Defokus von $D_A = 0.1$ *dpt* mit einem 10%-Fehler, also mit einer absoluten Genauigkeit von 0.01 *dpt*, eine Auflösung der Verschiebungs-Detektion von $\Delta x_d = 3.4$ μm erfordern würde. Diese Auflösung ist aufgrund einer optoelektronisch sinnvollen Mindest-Größe der realisierbaren Photodetektoren eines ASIC nicht möglich. Die Detektoreinheit des ASIC wird derart realisiert, dass die Detektion einer Aberration in einem dynamischen Bereich von $D_A = \pm 1$ *dpt* mit einer absoluten Auflösung von $\Delta D_A \approx 0.1$ *dpt* möglich ist.

Abbildung 1.9 stellt die Benutzeroberfläche eines Detektor-Simulators dar. Anhand des Simulators lassen sich verschiedene Aberrations-Typen konstruieren und das Muster der Fokalfunkte-Verschiebung unter Angabe von Fokallänge des Linsenarray, Apertur der Linsen und Anzahl der Linsen generieren. Die Qualität der Berechnung der Zernike-Koeffizienten lässt sich durch einen Least-Square-Fit unter Berücksichtigung der verschiedenen Auflösungen der Verschiebungs-Detektion abschätzen. Teilbild a) verdeutlicht die aufgrund einer vorgegebenen Wellenfront-Aberration entstehende Fokalfunkt-Verschiebung der 16×16 Fokalfunkte innerhalb der 256 Subaperturen. Die Größe der Detektor-Subaperturen, im Folgenden als *Cluster* bezeichnet, ist

³Diese Beziehung erfordert eine Umrechnung des auf μm bezogenen Koeffizienten C_4 in $\frac{1}{m}$ durch $C_4^* = \frac{1}{r_A^2} \cdot C_4$, mit r_A als Radius der Apertur, auf die C_4 bezogen ist.

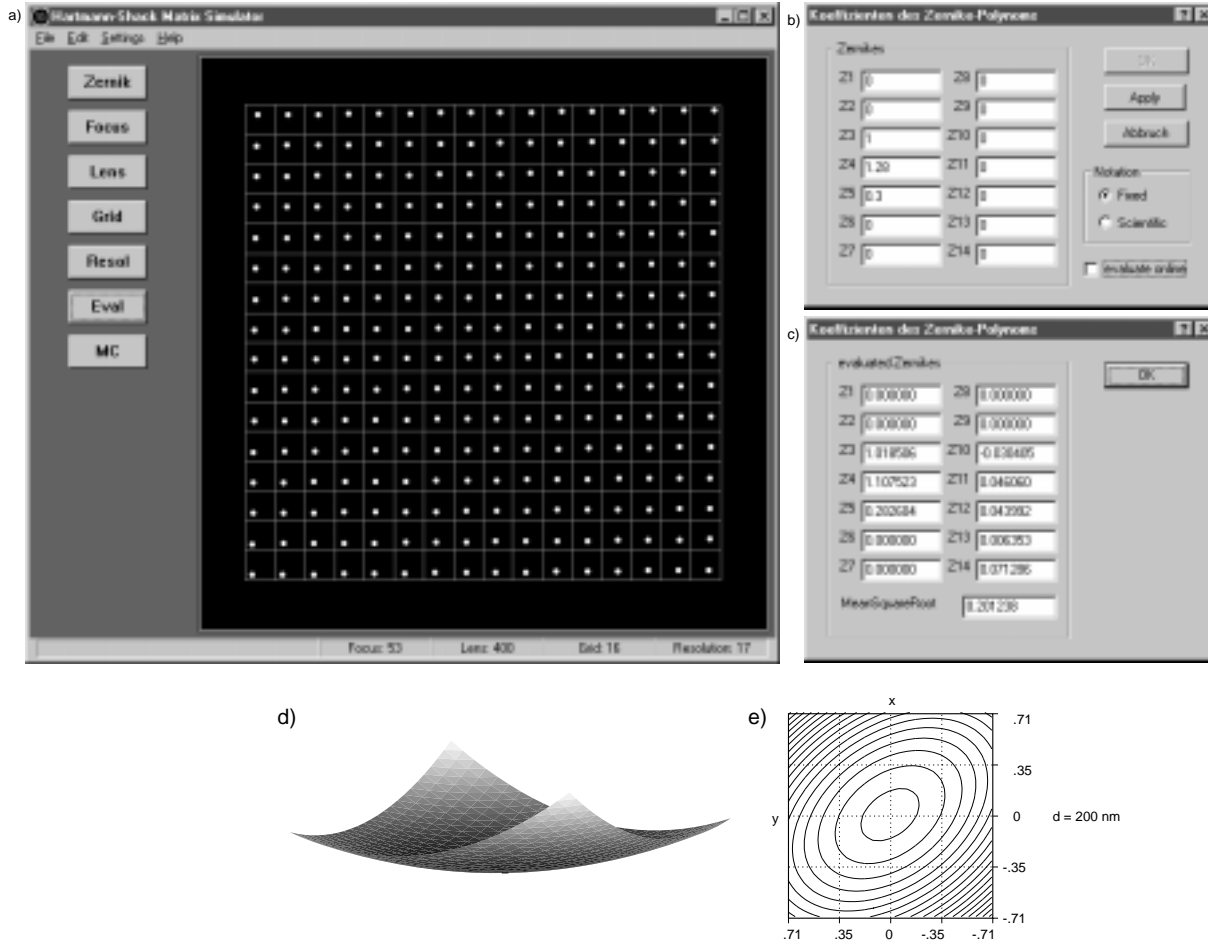


Abbildung 1.9: Simulator der Detektions-Auflösung des Wellenfrontsensor-ASIC

auf den Durchmesser $D_L = 400 \mu\text{m}$ der Linsen des Linsenarray abgestimmt; die Kantenlänge eines Clusters beträgt $400 \mu\text{m}$. Die Größe der abgebildeten Fokalfunkte ist maßstabsgerecht. Die Werte der 14 Zernike-Koeffizienten lassen sich wie in Teilbild b) gezeigt einstellen (vergleiche Tabelle 1.1). Zur Erzeugung des Musters aus a) wurden die Koeffizienten $C_3 = 1 \mu$, $C_4 = 1.28 \mu\text{m}$ und $C_5 = 0.3 \mu\text{m}$ gewählt, die praktischen Werten einer Hyperopie von 0.5 dpt und eines Astigmatismus von 0.4 dpt in 36° -Richtung entsprechen. Die Apertur der simulierten Wellenfront beträgt 6.4 mm . In den Teilbildern d) und e) ist der dreidimensionale Körper und das Höhenlinienmodell der Wellenfront dargestellt.

Teilbild c) zeigt die errechneten Zernike-Koeffizienten des Least-Square-Fit bei einer eingestellten Auflösung der Verschiebungs-Detektion von $\Delta x_d = 17 \mu\text{m}$. Der Algorithmus unterteilt dazu die lateralen Abweichungen der Fokalfunkte in diskrete Werte, die der gewählten Auflösung von $\Delta x_d = 17 \mu\text{m}$ entsprechen und ahmt auf diese Weise das tatsächliche Verhalten der Detektions-Matrix nach. Die Übereinstimmung der errechneten Koeffizienten mit den in Teilbild b) definierten Werten ist ein Maß der Least-Square-Fit-Qualität. Dazu ist zusätzlich die Standardabweichung (in c) *MeanSquareRoot*) der Koeffizienten angezeigt, um ein objektives Kriterium der Qualitätsbeurteilung zu geben. In dem dargestellten Fall ist eine absolute Genauigkeit der Bestimmung von C_4 von z.B. $\Delta C_4 = 0.172 \mu\text{m}$ möglich. Die Standardabweichung der errechneten Werte beträgt nach Teilbild c) $0.201 \mu\text{m}$ oder 0.078 dpt .

Aufgrund dieser Betrachtungen und Annahmen wird eine Auflösung der Detektion der Fokalfunk-

te-Verschiebung von $\Delta x_d = 15 - 20 \mu m$ als sinnvoll erachtet, um in einem dynamischen Bereich der Aberrationen von *ca.* $\pm 1.0 \text{ dpt}$ mit einer absoluten Genauigkeit von weniger als $\pm 0.05 \text{ dpt}$ messen zu können.

1.6.2 Diskriminierungsgenauigkeit der Verschiebungsdetektion

Da die Verteilung der Fokalkpunkt-Intensität einer Gauß'schen Verteilung ähnelt, existiert genau eine Stelle x^* mit $I(x^*) > I(x)$, $\forall x \in (-200 \mu m, +200 \mu m)$, $x \neq x^*$. Ein Detektionsalgorithmus zur Lokalisierung eines maximalen Wertes unterliegt einer bestimmten Diskriminierungsgenauigkeit, mit der dieses Maximum eindeutig bestimmt werden kann. Nach Definition der Detektorauflösung von $\Delta x_d = 15 - 20 \mu m$ erfolgt daher eine Betrachtung der Intensitätsverteilung innerhalb benachbarter x-Achsen-Abschnitte mit der Breite von Δx_d , um den relativen Intensitäts-Unterschied zwischen diesen zwei benachbarten Abschnitten zu bestimmen. Die Intensitätsverteilung eines Fokalkpunktes über x folgt nach [Bergmann 92] der Funktion

$$\begin{aligned} J(x) &= 4 J_0 \frac{J_1^2(z)}{z^2} \\ z &= \frac{2\pi}{\lambda} \cdot \frac{D_L}{f} \cdot x, \end{aligned} \quad (1.12)$$

mit $J_1(z)$ als Besselfunktion erster Art erster Ordnung, dem Linsen-Durchmesser D_L , der Fokallänge f und der Wellenlänge λ . Die Funktion $J(x)$ besitzt an der Stelle $x = \pm 63 \mu m$ die Minima 1. Ordnung; dieser Wert bestimmt den Durchmesser des *Airy-Scheibchens*. Zur Berechnung der Intensitätsverteilung über einer bestimmten lateralen Fläche wird die Gleichung 1.12 zweidimensional entwickelt:

$$\begin{aligned} J(x, y) &= 4 J_0 \frac{J_1^2(\sqrt{z_x^2 + z_y^2})}{z_x^2 + z_y^2} \\ z_x &= \frac{2\pi}{\lambda} \cdot \frac{D_L}{f} \cdot x \\ z_y &= \frac{2\pi}{\lambda} \cdot \frac{D_L}{f} \cdot y \end{aligned} \quad (1.13)$$

Diese Funktion ist in Abbildung 1.10 a) in den Grenzen von $x, y \in (-200 \mu m, +200 \mu m)$, normiert auf die maximale Intensität, dargestellt. Teilbild b) zeigt einen vergrößerten Ausschnitt des Fokalkpunktes in den Grenzen von $x, y \in (-60 \mu m, +60 \mu m)$. Das geometrische Mittel der

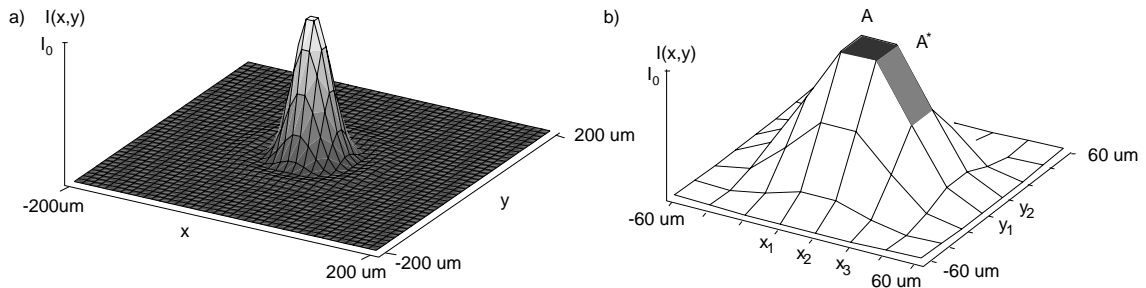


Abbildung 1.10: Intensitätsverteilung eines Fokalkpunktes nach Airy

Intensität einer Fläche A innerhalb x_1, x_2 und y_1, y_2 wird durch Integration von Gl. 1.13 bestimmt:

$$J_A = \frac{4 J_0}{A} \cdot \int_x \int_y Z(x, y) dy dx$$

$$Z(x, y) = \frac{J_1^2(\sqrt{z_x^2 + z_y^2})}{z_x^2 + z_y^2} \quad (1.14)$$

Gleichung 1.14 wird numerisch für zwei benachbarte Flächen A und A^* mit $x_1 = -\frac{1}{2}\Delta x_d$, $x_2 = +\frac{1}{2}\Delta x_d$, $x_3 = +\frac{3}{2}\Delta x_d$, $y_1 = -\frac{1}{2}\Delta x_d$ und $y_2 = +\frac{1}{2}\Delta x_d$ berechnet. Unter Annahme von $f = 53 \text{ mm}$, $D_L = 400 \mu\text{m}$, $\Delta x_d = 17 \mu\text{m}$ und $\lambda = 780 \text{ nm}$ folgt ein Verhältnis von

$$\frac{J_A^*}{J_A} = \frac{\frac{4 J_0}{A^*} \cdot \int_{x_2}^{x_3} \int_{y_1}^{y_2} Z(x, y) dy dx}{\frac{4 J_0}{A} \cdot \int_{x_1}^{x_2} \int_{y_1}^{y_2} Z(x, y) dy dx} = \frac{\int_{x_2}^{x_3} \int_{y_1}^{y_2} Z(x, y) dy dx}{\int_{x_1}^{x_2} \int_{y_1}^{y_2} Z(x, y) dy dx} = 0.767. \quad (1.15)$$

Die Genauigkeit der Diskriminierung zwischen zwei benachbarten Intensitäten muss folglich mindestens 23.2% betragen.

Die Form des Fokalfunktes gleicht der oben besprochenen Funktion 1.12 jedoch nur, wenn die Wellenfront innerhalb der Apertur der Linse selbst keine Aberrationen besitzt. Dieser Fall ist nicht zu ermöglichen. Die Abschätzung der Diskriminierungs-Qualität wird folglich um die Erkenntnisse aus Messungen der realen Fokalfunkt-Gestalt erweitert. In Abbildung 1.11 a) ist die Intensitätsverteilung eines mit einer CCD-Kamera erfassten Fokalfunktes dargestellt. Die

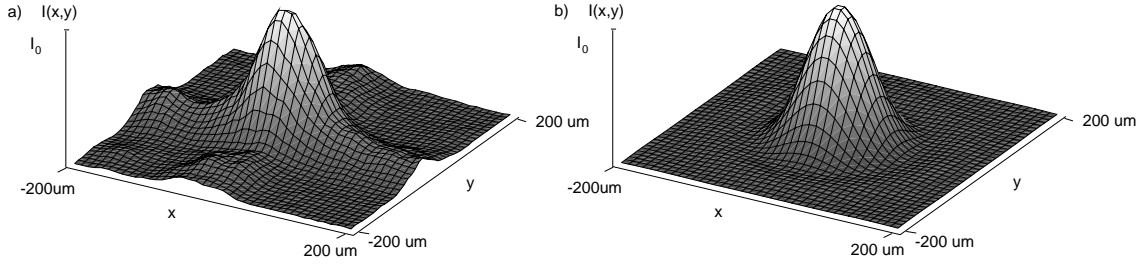


Abbildung 1.11: Gemessene und angenäherte Intensitätsverteilung eines Fokalfunktes

gemessenen Intensitätsverteilung in einer Umgebung von $\frac{3}{2}\Delta x_d$ um das Maximum kann durch Annahme eines geringeren Linsen-Durchmessers mit $D_L = 225 \mu\text{m}$ hinreichend genau simuliert werden, vergleiche Abbildung 1.11 b). Unter Anwendung von Gl. 1.15 und Wahl des Linsen-Durchmessers von $D_L = 225 \mu\text{m}$ erfolgt ein geringeres Intensitätsverhältnis von $\frac{J_A^*}{J_A} = 0.92$. Eine Diskriminierungs-Genauigkeit von 8% ist notwendig.

1.6.3 Zeitliche Dauer der Detektion

Die Detektion der Fokalfunkte-Verschiebungen und die Berechnung der Zernike-Koeffizienten einer Wellenfront soll mit einer Wiederholrate von 1 kHz ermöglicht werden. Dies entspricht einer verfügbaren Zeitdauer von 1 ms für eine Wellenfrontbestimmung. Die Frequenz von 1 kHz ist zum einen erforderlich, um die in Kapitel 1.3 beschriebene Iteration der Aberrations-Kompensation in einer für den Patienten zumutbaren Zeitdauer zu leisten, zum anderen, um

die durch die inhärenten Augenbewegungen des Patienten entstehende mögliche Unschärfe der Aberrations-Detektion zu vermeiden. Die Augenbewegungen erfolgen mit einer Frequenz von ca. 80Hz und verursachen eine Verschiebung der Retina im Fokus des abtastenden Laserstrahls um 2-3 Rezeptoren (ca. $10\text{ }\mu\text{m}$) [Carpenter 88]. Durch diese Augenbewegungen wird ein Ausbleichen der Photorezeptoren der Retina verhindert.

1.6.4 Photosensitivität der Detektion

Nach den Unfallverhütungsvorschriften im Umgang mit Lasern ist für eine Zeitdauer von weniger als 30 Sekunden eine direkte Einstrahlung in das Auge mit einer Dauerleistung von maximal $140\text{ }\mu\text{W}$ zulässig. Zur Sicherheit wird eine Leistung von $P_L \approx 100\text{ }\mu\text{W}$ gewählt. Bei einer Wellenlänge $\lambda = 632\text{ nm}$ wird ein Anteil von ca. 1% diffus reflektiert und gelangt in das System des Abtasters. Unter Berücksichtigung der verschiedenen optischen Elemente im Strahlengang des Systems (vergleiche Abbildung 1.4) kann auf eine direkt vor der Wellenfrontdetektions-Einheit einfallende Lichtleistung von ca. 50-500 nW geschlossen werden. Diese Lichtleistung wird durch die Linsen in Fokalfunkten konzentriert. Bei Annahme einer Beleuchtung aller Linsen der 16×16 -Matrix ergibt sich eine Intensität eines Fokalfunktes von $J_{FP} \approx 200\text{ pW/Spot}$ - 2 nW/Spot . Unter Annahme eines Fokalfunkt-Durchmessers von $D_{FP} \approx 100\text{ }\mu\text{m}$ und einer homogenen Intensitätsverteilung des Spot entspricht dies einer zur Photonen-Detektion auf der Chip-Oberfläche verwendbaren Lichtintensität von $J_{FP} \approx 0.025 - 0.25\text{ W/m}^2$.

1.6.5 Zusammenfassung der Spezifikationen

Die Entwicklung des Wellenfrontsensor-ASIC geschieht unter Zielsetzung der folgenden Spezifikationen:

- Die Auflösung der Verschiebungs-Detektion der Fokalfunkte beträgt $\Delta x_d \approx 15 - 20\text{ }\mu\text{m}$.
- Der dynamische Bereich der Verschiebungs-Detektion der Fokalfunkte ist durch die Durchmesser der Linsen des verwendeten Array festgelegt als $\Delta x_{max} = \pm 200\text{ }\mu\text{m}$.
- Die Genauigkeit der Intensitäts-Diskriminierung zur Bestimmung der Fokalfunkt-Position muss besser als 8% sein.
- Die Wiederholrate der Wellenfront-Detektion und -Berechnung ist 1 kHz .
- Die zur Verfügung stehende Lichtintensität pro Fokalfunkt beträgt mindestens $J_{FP} \approx 200\text{ pW/Spot}$

Kapitel 2

Die optoelektronischen Möglichkeiten eines ASIC

Zu Beginn des Kapitels erfolgt eine Erläuterung der zur Herstellung des ASIC verwendeten Prozesstechnologien. Verschiedene in diesen Technologien realisierbare optische Sensoren werden im Hinblick auf spektrale Empfindlichkeit und Geschwindigkeit beschrieben, um die Motivation zu ihrer Verwendung als Photodetektoren im ASIC zu verdeutlichen. Weiterhin findet eine Betrachtung der Subthreshold-Eigenschaften von MOS-Transistoren statt, um die Qualität der Verarbeitung geringer Photoströme abzuschätzen.

Die Realisierung des Wellenfrontdetektors durch einen applikationsspezifischen integrierten Schaltkreis ist notwendig, um die Projekt-Spezifikation der Detektionsrate von 1 kHz zu erfüllen. Die im Allgemeinen eingesetzten Wellenfront-Detektorsysteme verwenden als bildgebendes Medium eine CCD-Kamera und führen die Bildinformationen des Fokalpunkte-Musters einem Datenverarbeitungssystem zu. Die Bestimmung der lateralen Verschiebung der Fokalpunkte unterliegt einem Software-Algorithmus. Die Detektionsrate des Systems bestimmt sich aus der Bildaufnahme-Frequenz der CCD-Kamera und aus der Verarbeitungsgeschwindigkeit der Datenverarbeitung. Die typische Video-Frequenz einer CCD-Kamera von 50 Hz definiert zumeist eine obere Grenzfrequenz der Detektionsrate des Systems. Unter Anwendung einer modifizierten CCD-Kamera mit höheren Bildauslese-Frequenzen kann diese Grenze überschritten werden, wobei die Datenverarbeitung zur Bestimmung der Fokalkpunkt-Verschiebung die entstehende erhöhte Datenrate zu verarbeiten hat.

Die Fähigkeit des Systems zur Echtzeitverarbeitung, also die Möglichkeit zur Verarbeitung der Informationen mit der gleichen Frequenz der Bildaufnahme, ist ein Maß zur Bestimmung der Leistungsfähigkeit. Die Lösung dieser Problematik ist Gegenstand der Forschung; es bestehen gegenwärtig hauptsächlich zwei Möglichkeiten zur Realisierung eines Echtzeit-Systems. Zum einen ist es möglich, parallelisierte Rechenoperationen in speziellen Prozessoren auszuführen. Diese Schematik führt von einfachen digitalen Signalprozessoren (DSPs) über hochspezialisierte Prozessoren mit durch Hardware implementierter Funktionen bis hin zu der gleichzeitigen Anwendung mehrerer dieser Prozessortypen. Zum anderen werden Systeme entwickelt, die durch

den Aufbau mit programmierbaren Gatter- oder Operator-Bausteinen (FPGAs¹, FPOAs²) eine massive Parallelisierung des Algorithmus erreichen und Echtzeitbildverarbeitung ermöglichen.

Die Entwicklung eines ASIC als Wellenfrontdetektor wird motiviert durch die Möglichkeit, durch eine einzige Systemkomponente optische Signale in elektrische zu konvertieren und weiterzuverarbeiten; die Kombination aus Opto-Sensorik und Signalverarbeitung ermöglicht eine hohe Signalverarbeitungs-Geschwindigkeit. Der Wellenfrontdetektor-ASIC gehört daher zu einer Klasse von integrierten Schaltkreisen, die in der Literatur als *Vision-Chips* bezeichnet werden; der Begriff beschreibt die Möglichkeit der Bildaufnahme- und Verarbeitung in *einem* integrierten Schaltkreis.

2.1 Die Prozesstechnologien

Die Entwicklung der Prototypen des Wellenfrontsensor-ASIC fand unter Anwendung des $0.8\mu\text{m}$ -CMOS³-, des $0.6\mu\text{m}$ -CMOS- und des $0.8\mu\text{m}$ -BiCMOS⁴-Prozesses der Firma *Austria Mikro Systeme International* (AMS) statt. Zum Verständnis der weiteren Ausführungen dieses Kapitels ist in Abbildung 2.1 ein Querschnitt der beiden Prozess-Familien CMOS und BiCMOS gezeichnet, in dem die aktiven Elemente, die in diesen Prozessen implementiert werden können, dargestellt sind. Die Strukturen, die eine vom Substrat verschiedene Konzentration von Dotierstoffen besitzen, sind durch Graustufen hervorgehoben. Die an den Strukturen angesetzten Abkürzungen n -, $n+$, $n++$ und p -, $p+$, $p++$ symbolisieren den Typ und die Höhe der Konzentration der Fremdatome in dem entsprechenden Gebiet. Mit Hilfe der Bezeichnungen B , D , G , S , K , A , KO , E sind die Funktionen der Strukturen als Bulk, Drain, Gate, Source, Kathode, Anode, Kollektor und Emitter gekennzeichnet. Die Angabe (Gnd) bedeutet, dass das Potential der Struktur aufgrund der Regeln des Schaltungsentwurfs von CMOS-Schaltungen dem niedrigsten Potential der Schaltung entspricht; dies ist im Allgemeinen das Masse-Potential *Ground*. Teilbild a) stellt

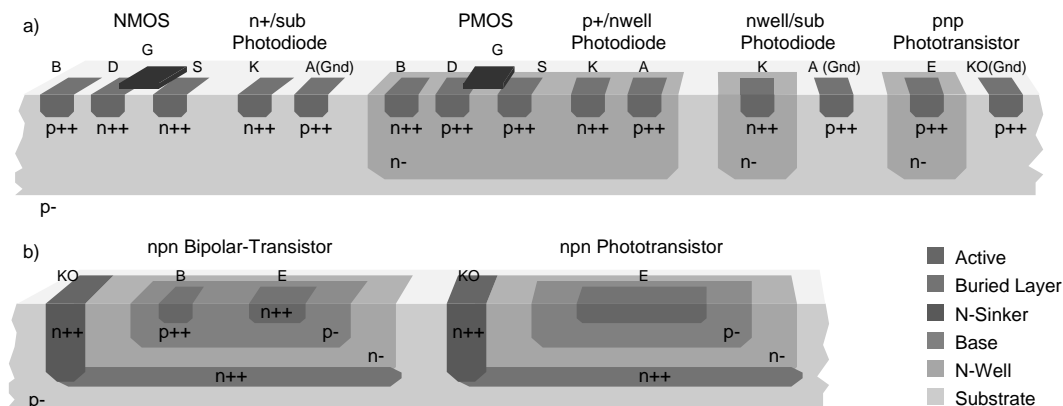


Abbildung 2.1: Querschnitt eines Wafers des CMOS- und BiCMOS-Prozesses

den Querschnitt des in CMOS-Technologie gefertigten Wafers dar. Die beiden komplementären MOS-FET-Strukturen, sowie die drei realisierbaren Diodentypen (Photodiode) sind dargestellt. Zusätzlich ist ein pnp-Transistor, der ebenfalls zur Photogeneration verwendet werden kann,

¹Field Programmable Gate Arrays

²Field Programmable Operator Arrays

³Complementary Metal Oxide Silicon

⁴Bipolar-CMOS

abgebildet. Dieser Transistor wird im Allgemeinen als *parasitärer* vertikaler npn-Transistor bezeichnet, da sein bei Verwendung einer N-Wanne zwangsläufiges Auftreten durch unerwünschte Verstärkungsfaktoren zu Latch-Up-Effekten führen kann [Allen 87]; er wird jedoch in vielen Applikationen als Phototransistor eingesetzt. In Teilbild b) ist ein Ausschnitt des in BiCMOS-Technologie gefertigten Wafers gezeigt. Dieser Prozess ist eine Erweiterung des oben angeführten CMOS-Prozesses; es sind alle vorangehend beschriebenen Elemente implementierbar. Als spezielle Struktur ist ein vertikaler npn-Transistor entwickelt worden, der eine hohe Stromverstärkung bietet; dieser Transistor besitzt als Phototransistor eine erhöhte spektrale Empfindlichkeit. Im Folgenden werden die fundamentalen Eigenschaften der Photodioden- und Transistoren, sowie die der beiden NMOS- und PMOS-Transistoren erläutert.

2.2 Photodetektoren

Der Photoeffekt beschreibt die Änderung der elektrischen Eigenschaften eines Halbleiters durch Absorption elektromagnetischer Strahlung. Generell gibt es drei Formen der Wechselwirkung elektromagnetischer Strahlung mit Festkörpern: Compton-Streuung, Paarbildung und Photoeffekt. Es werden nur die Vorgänge des Photoeffektes bei Photonen-Energien von $W_{ph} \in (1.15 \text{ eV}, 4 \text{ eV})$ betrachtet, dem Bereich der sichtbaren Strahlung von nahem IR (Infrarot) bis zu nahem UV (Ultraviolett). Kristallines Silizium, aus dem integrierte Schaltkreise hergestellt werden, besitzt in diesem Energiebereich eine starke Absorption von Photonen und ist daher für photosensorische Anwendungen geeignet. Es gibt verschiedene Arten von Photodetektoren, die sich hauptsächlich in der Verwertung der durch Absorption erzeugten Ladungsträger unterscheiden [Bergmann 92]. Im Folgenden werden nur die *ladungstrennenden* Photodetektoren wie Photodioden oder Phototransistoren behandelt.

Ein Maß der spektralen Empfindlichkeit von Photodetektoren ist die *Quanteneffizienz*. Sie beschreibt das reziproke Verhältnis der Menge der einfallenden Photonen in einer Zeiteinheit zu der Menge der durch Absorption als Photostrom nutzbaren Ladungsträger-Paare.

$$\eta = \frac{I_{ph}/e}{P_{ph}/W_{ph}} \quad (2.1)$$

P_{ph} beschreibt die Leistung der elektromagnetischen Strahlung, W_{ph} die Energie eines Photons der Strahlung, I_{ph} den Photostrom und e die Elementarladung. Aus der Quanteneffizienz lässt sich der Photostrom berechnen, der durch Trennung der absorbierten Ladungsträger erzeugt werden kann:

$$I_{ph} = \frac{e \eta P_{ph} \lambda}{h c} \quad (2.2)$$

Dabei ist λ die Wellenlänge der Strahlung, h das Plank'sche Wirkungsquantum und c die Lichtgeschwindigkeit. Die Quanteneffizienz der Photonengeneration ist abhängig von dem Absorptionsverhalten des Detektormaterials und der Effizienz der Trennung der absorbierten Ladungsträger. Diese beiden Eigenschaften werden im Folgenden näher betrachtet.

2.2.1 Strahlungsabsorption

Der physikalische Effekt der Absorption in einem Halbleiterkristall beruht auf der Umwandlung der Energie eines Photons in Phononen⁵ oder Ladungsträger-Generation, also Anhebung von Elektronen in das Leitungsband-Energieniveau. Zur Berechnung der Intensität $J_{ph}(z)$ in

⁵Gitterschwingungen

Richtung des k -Vektors einer einwirkenden elektromagnetischen Strahlung sei zuerst der Anteil J_{pht} bestimmt, der nach Reflexion eines bestimmten Anteils in die Oberfläche eindringt. Unter Annahme senkrechter Bestrahlung bestimmt sich das Verhältnis von transmittierter zu eingestrahelter Intensität wie

$$J_{\text{pht}} = J_{\text{ph0}}(1 - \sigma) \quad (2.3)$$

mit σ als Reflektivität der Grenzschicht. Unter Aufbringen einer Antireflex-Beschichtung lässt sich der Anteil der reflektierten Intensität verringern. Nach [Jones 87] ergibt sich für Si z.B. eine aufzubringende Dicke von SiO_2 von $d \approx 100 \text{ nm}$.

Der transmittierte Anteil der Intensität sei J_{ph} . Die Absorption der Photonen von J_{ph} erstreckt sich über eine bestimmte Strecke, innerhalb derer die Abnahme der Intensität einem exponentiellen Verlauf gleicht. Auf der Wegstrecke Δz wird ein konstanter Anteil der Photonen $\alpha \cdot \Delta J$ absorbiert. Es folgt aus dem differentiellen Zusammenhang $\frac{dJ_{\text{ph}}}{dz} = -\alpha \cdot J_{\text{ph}}(z)$ das Lambert'sche Gesetz der Absorption:

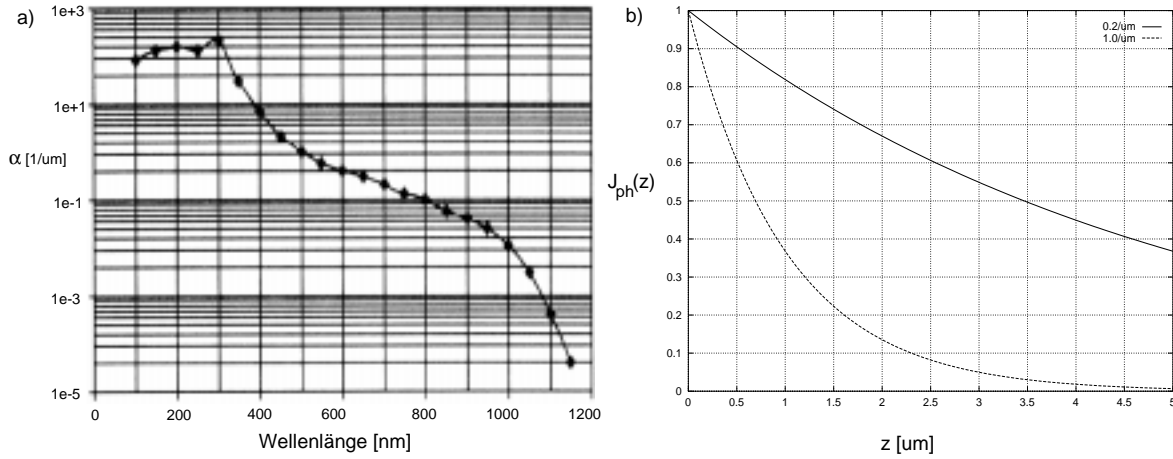
$$J_{\text{ph}}(z) = J_{\text{ph0}} \cdot e^{-\alpha z} \quad (2.4)$$

Der Absorptions-Koeffizient α eines Festkörpers ist abhängig von der Energie der eindringenden Photonen. Unterhalb der Energie, die der Differenz zwischen Valenzband-Energieniveau und Leitungsband-Energieniveau, also der Bandlücken-Energie E_g entspricht, besitzen die Photonen nicht genügend Energie, um ein Elektron in das Leitungsband zu heben. Die dennoch in geringem Maße entstehende Absorption findet ihre Ursache in der Existenz von Defekt-Energie-Zuständen in der Bandlücke, in der Absorption von freien Elektronen, also Elektronen, die sich schon im Leitungsband befinden, und in einer Anzahl von Energiezuständen dicht unterhalb der Leitungsbandkante, die durch die zufällige Donatorverteilung entstehen. Ab $W_{\text{ph}} > E_g$ vergrößert sich α um einige Größenordnungen, da die Anzahl der besetzbaren Energiezustände als Funktion der Energie steigt. Gleichung 2.5 zeigt die Anzahl der Energie-Zustände, die von einem Elektron in einkristallinem Silizium besetzt werden können [Sze 85]:

$$n_n(E) = 4\pi \left(\frac{2m_n^*}{h^2} \right)^{\frac{3}{2}} \sqrt{E - E_g} \quad (2.5)$$

Dabei ist m_n^* die effektive Masse eines Elektrons und E_g die Energie-Differenz zwischen Valenzband- und Leitungsband-Energieniveau. Je größer die Anzahl der von einem Elektron erreichbaren Zustände ist, um so wahrscheinlicher ist die Absorption. In Silizium ist zusätzlich zu der Energie, die erforderlich ist, ein freies Elektron zu erzeugen, noch eine Impulsänderung des angeregten Elektrons notwendig, um in dicht oberhalb der Leitungsbandkante liegende Energieniveaus zu gelangen. Dies ist eine Ursache der unterschiedlichen k -Vektoren von Valenzband-Maximum und Leitungsband-Minimum. Ein Übergang eines Elektrons kann nur unter gleichzeitigem Austausch von Energie mit dem Kristallgitter in Form von Phononen stattfinden. Dies ist eine Eigenheit der *indirekten* Halbleiter, wie zum Beispiel Silizium oder Germanium. Ab einer bestimmten Energie ist die Besetzung eines Energie-Zustandes ohne Änderung des Impulses möglich. Bei geringeren Wellenlängen resultiert ein starker Anstieg von α . UV-Strahlung wird daher direkt unter der Oberfläche des Wafers absorbiert und kann im Allgemeinen nicht durch Silizium-Photodetektoren erfasst werden. Die Abhängigkeit α von λ ist für Silizium in Abbildung 2.2 a) dargestellt [Wolf 71]. In Teilbild b) ist der Verlauf der Intensität im Halbleiter für die Werte $\alpha = 0.2/\mu\text{m}$ und $\alpha = 1.0/\mu\text{m}$ über der Tiefe im Halbleiter dargestellt.

Zur Nutzung der erzeugten Ladungsträger ist in dem Falle der Photostromgeneration eine *Trennung* der beiden Ladungsträger-Arten notwendig. Diese Trennung erfolgt in für den Photodetektor charakteristischen geometrischen Bereichen; dichter unter der Oberfläche liegende Detektoren können bei $\alpha = 1.0/\mu\text{m}$ beispielsweise effizienter arbeiten als tiefer liegende. Durch Integration

Abbildung 2.2: Absorptions-Koeffizient α und Intensitätsverlauf in Silizium

über einen Bereich Δz ist die Anzahl n der in dem Volumen $\Delta z \cdot A$ entstehenden Ladungsträger-Paare pro Zeiteinheit bestimmbar, und damit der Photostrom, der durch vollständige Ladungstrennung generierbar ist:

$$I = e \cdot n = e \cdot \frac{A \lambda}{h c} \int_{z_0}^{z_0 + \Delta z} \frac{dJ_{ph}(z)}{dz} dz \quad (2.6)$$

Effiziente Detektoren verfügen über einen großen Wert von Δz und einen kleinen Wert von z_0 .

2.2.2 Photodioden

Um die Definition einer Quanteneffizienz in Photodioden zu erläutern, erfolgt an dieser Stelle eine kurze Beschreibung der physikalischen und elektrischen Eigenschaften von Halbleiterdioden. Dioden repräsentieren die physikalische Verbindung zweier Schichten mit gegensätzlichen Leitfähigkeitstypen. Die eine Schicht, im Folgenden als n-Schicht bezeichnet, wird im Herstellungsprozess durch Einbringen von 5-wertigen Donator-Atomen wie Phosphor oder Arsen erzeugt. Die zweite Schicht, im Folgenden als p-Schicht bezeichnet, wird durch Einbringen eines 3-wertigen Akzeptor-Atoms wie Bor erzeugt. Die Konzentration, mit der die jeweilige Dotierung stattfindet, sowie die Gestalt des Konzentrationsprofils bestimmen die elektrischen Eigenschaften.

Für die Untersuchung der Quanteneffizienz wird speziell die Grenzschicht des Überganges betrachtet. Im thermischen Gleichgewicht entsteht eine Verarmungszone (Raumladungszone) zwischen den beiden Schichten. Der Konzentrationsgradient repräsentiert einen Unterschied in der in beiden Schichten vorhandenen Menge von Ladungsträgern der gleichen Sorte. Dadurch entsteht ein Diffusionsstrom, der zur Einhaltung des thermischen Gleichgewichts durch einen gegenläufigen Driftstrom ausgeglichen werden muss. Es resultiert ein elektrisches Feld, dessen Feldstärke von dem Konzentrationsprofil der Dotierstoffe abhängig ist. Durch dieses Feld werden die durch Absorption entstehende Ladungsträger getrennt und es kann ein Photostrom generiert werden. Die Photoeffizienz einer Photodiode ist daher in erster Näherung eine Funktion der Breite W ihrer Verarmungszone und dem Abstand der Verarmungszone von der Oberfläche, entsprechend Gleichung 2.6. Die Breite der Raumladungszone W eines abrupten pn-Überganges im thermischen Gleichgewicht, unter Anlegen einer Spannung V , wird durch die folgende

Beziehung angenähert [Sze 85]:

$$W = \sqrt{\frac{2\epsilon_{si}}{e} \frac{N_A + N_D}{N_A \cdot N_D} (V_D - V)} \quad \text{mit} \quad (2.7)$$

$$V_D = \frac{kT}{e} \ln \frac{N_A N_D}{n_i^2} \quad (2.8)$$

Dabei ist k die Boltzmann-Konstante, e die Elementarladung, ϵ_{si} die Dielektrizitätskonstante von Silizium, N_A und N_D die Konzentration der Akzeptor und Donatoratome und n_i die intrinsische Ladungsträger-Konzentration bei Raumtemperatur. In dieser Beziehung wird eine abrupte Konzentrationsänderung mit einer definierten Grenze zwischen der n- und der p-Schicht angenommen. Tatsächlich existiert jedoch ein Konzentrationsgradient, der von der vertikalen Tiefe im Halbleiter abhängig ist; die Weite der Raumladungszone wird daher mit Gleichung 2.8 ungenau quantifiziert. Der Gradient der Konzentration entsteht nach Einbringen der Fremdatome mit den verschiedenen Mechanismen der Implantation oder Diffusion und verändert sich durch thermische Diffusion der Dotierstoff-Atome während der aufeinanderfolgenden Prozessschritte der Herstellung; im Allgemeinen ist der Gradient nicht durch eine Funktion erster Ordnung zu beschreiben. Abbildung 2.3 stellt den Verlauf der Konzentration der drei im CMOS/BiCMOS-Prozess nutzbaren pn-Übergänge als Funktion der Tiefe im Halbleiter dar. Die abgebildeten

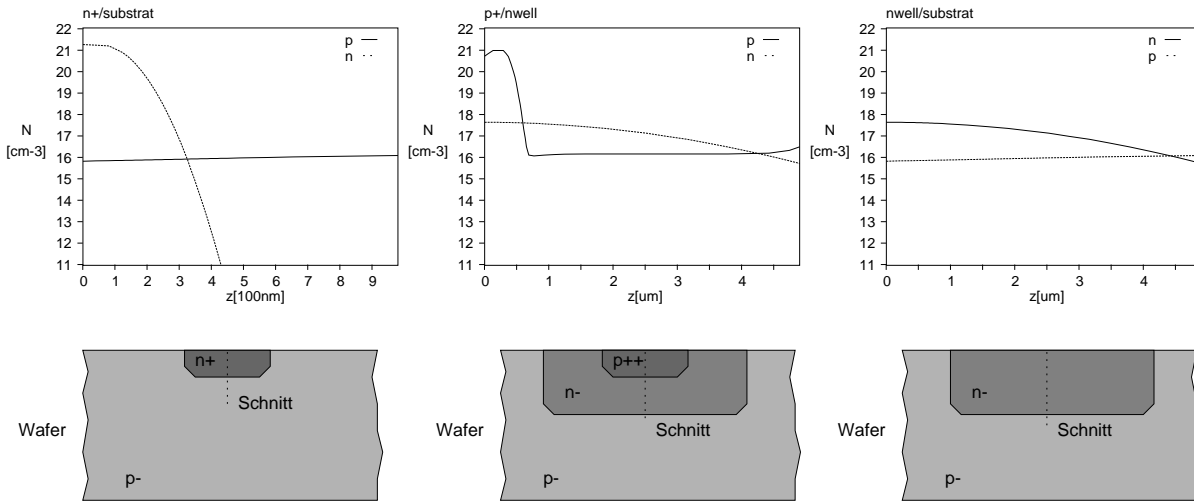


Abbildung 2.3: Dotierprofile der im CMOS-Prozess verwendbaren Photodioden

Profile entstammen Prozess-Simulations-Daten der Firma *AMS* und zeigen die tatsächlich im Prozess realisierten Konzentrations-Profile. Die Profile wurden zur Evaluierung der spektralen Empfindlichkeiten eingesetzt und in [Schneider 98] als Grundlage einer Device-Modellierung mit dem Device-Simulator *ATLAS* der Firma *Silvaco International* verwendet. Die Simulation der Quanteneffizienz erfolgt durch den Device-Simulator unter Eingabe der Konzentrationsprofile und der Elektroden-Potentiale. Die spektrale Empfindlichkeit der drei möglichen Photodioden-Typen unter Anlegen einer Sperrspannung von $V = -2\text{ V}$ ist in Abbildung 2.4 zu sehen. In Teilbild a) ist der simulierte Verlauf der Quanteneffizienz der drei Photodioden der CMOS-Technologie dargestellt. Zusätzlich sind in [Loose 96] bestimmte Messwerte von η für die drei pn-Übergänge eingetragen. Der Abzissen-Abschnitt von $\lambda \approx 680\text{ nm}$ entspricht der Strahlungsfrequenz des für die Messung verwendeten Halbleiter-Diodenlasers. In Teilbild b) ist die in [Teodori 97] mit einem Spektrometer gemessene Quanteneffizienz einer n+/sub-Diode abgebildet. Die Interferenzen der einfallenden Strahlung in der Passivierungs- und Feldoxyd-Schicht

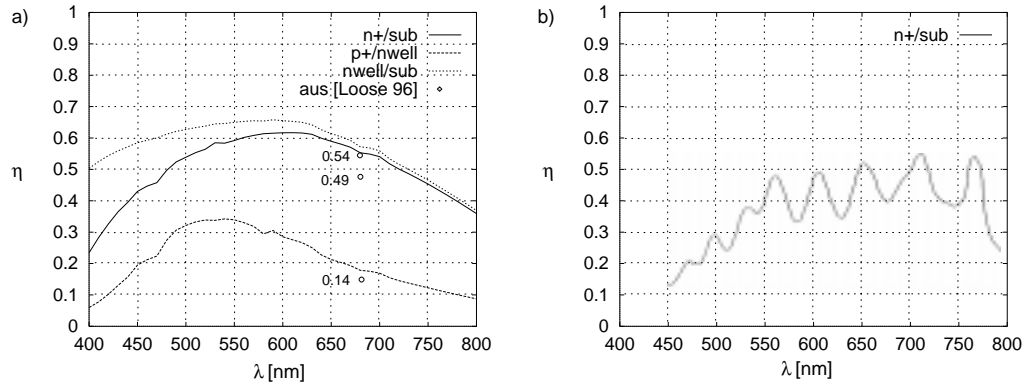


Abbildung 2.4: Quanteneffizienzen der Photodioden

an der Oberfläche des Wafers erzeugen niederfrequente Veränderungen der Quanteneffizienz als Funktion der Wellenlänge. Aus den beiden Schaubildern lässt sich ableiten, dass die spektrale Empfindlichkeit der n+/sub- und der nwell/sub-Diode, deren Raumladungszonene sich über eine große Tiefe erstrecken kann, höhere Werte besitzt. Die Raumladungszonene der p+/nwell-Diode ist durch den zweiten pn-Übergang der nwell/sub-Diode begrenzt, so dass nur wenige der absorbierten Photonen zu einem Photostrom beitragen können.

In [Schneider 98] wurde ebenfalls durch Simulation das zeitliche Verhalten der Photodioden bei Änderungen der Photonenintensitäten abgeschätzt. Dazu wurde eine Transientenanalyse des Photostromes bei sprunghörmiger Änderung der Intensität von $J_{ph} = 1 \text{ W/m}^2$ auf $J_{ph} = 0 \text{ W/m}^2$ durchgeführt. Tabelle 2.1 fasst die Ergebnisse dieser Simulationen zusammen; τ repräsentiert die Zeitdauer zur Annäherung des Endwertes des Photostromes auf 10% nach Abschalten des Photonenflusses. Es zeigt sich, dass die p+/nwell-Photodiode die geringste Zeitverzögerung in der Änderung des Photostromes besitzt. Dies ist eine Konsequenz der im Vergleich zu den anderen beiden Photodioden-Typen schmalen Raumladungszonene, in der weniger Ladungsträger rekombinieren müssen.

	n+/sub	p+/nwell	nwell/sub
$\tau \text{ [ns]}$	1.5	0.62	3.7

Tabelle 2.1: Sprungantwortzeit der Photodioden aus [Schneider 98]

Abbildung 2.5 zeigt das Ersatzschaltbild einer in Sperrrichtung betriebenen Photodiode. Zwei

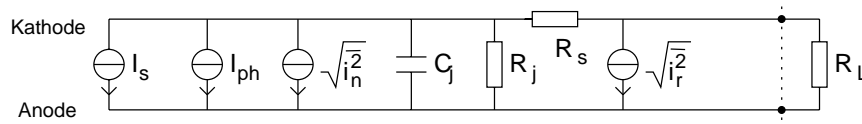


Abbildung 2.5: Ersatzschaltbild einer Photodiode

grundlegende Parameter der Diode werden durch den Sperr-Sättigungsstrom I_s und die Sperrschicht-Kapazität C_j modelliert. Der elektrische Widerstand der Sperrschicht R_j ist im Allgemeinen sehr hoch und kann daher zumeist vernachlässigt werden. Der Serienwiderstand R_s beschreibt die

Schichtwiderstände der Zuleitungszonen und Anschlüsse. A ist die Fläche der Sperrschicht, L_n und L_p sind die Diffusions-Längen, D_n und D_p die Diffusions-Konstanten der Minoritäten.

$$C_j = A \sqrt{\frac{\epsilon_{si} e N_A N_D}{2(N_A + N_D)(V_D - V)}} \quad (2.9)$$

$$I_s = e A n_i^2 \left(\frac{D_p}{L_p N_D} + \frac{D_n}{L_n N_A} \right) \quad (2.10)$$

$$R_j \rightarrow \infty \quad (2.11)$$

Der Photostrom, der unter Absorption von Photonen entsteht, wird durch die Stromquelle I_{ph} dargestellt und mit Gleichung 2.2 berechnet. Eine weitere charakteristische Eigenschaft der Photodiode ist der statistische Prozess der Photostrom-Generation und des thermisch induzierten Ladungsträger-Transports in der Verarmungszone. Die Ursache dieses *Rauschens* findet sich in der Zufälligkeit, mit der die Ladungsträger unter thermischer oder absorptionsunterstützter Energie die Potentialbarriere der Sperrschicht überwinden. Stromquelle $\sqrt{i_n^2}$ repräsentiert diesen als *Schrotrauschen* bezeichneten Prozess. Zusätzlich ist das allgemeine thermische Rauschen der Ladungsträger in Widerständen durch $\sqrt{i_r^2}$ einbezogen.

$$\overline{i_n^2} = 2 e (I_{ph} + I_s) \Delta f. \quad (2.12)$$

$$\overline{i_r^2} = 4 k T \frac{1}{R} \Delta f. \quad (2.13)$$

R ist eine Reihen-Parallel-Schaltung aus R_j , R_s und R_L . Δf gibt das Frequenzintervall an, in dem das Rauschen auf das elektrische Verhalten der Schaltung Einfluss hat. Die Bandbreite ist unter Berücksichtigung der Last-Impedanz abzuschätzen. Eine allgemeine Angabe des Rauschanteils pro Frequenzintervall wird durch den Ausdruck $i_n^2 = 2 e (I_{ph} + I_s)$ in Einheiten von $[\frac{A^2}{Hz}]$ gegeben. Eine wichtige Größe zur Definition der Rauscheigenschaften eines Systems ist das *Signal-Rausch-Verhältnis* $\frac{S}{N}$, das das Verhältnis des Nutzsignals zum Rauschen ausdrückt. Nach Abbildung 2.5 berechnet sich das Signal-Rausch-Verhältnis der in R_L umgesetzten Leistung zu

$$\frac{S}{N} = \frac{I_{ph}^2 R_j \|(R_L + R_s)}{(\overline{i_n^2} + \overline{i_r^2}) R_j \|(R_L + R_s)} \frac{1}{\Delta f} = \frac{(e \eta P_{ph} \lambda / h c)^2}{(2 e (e \eta P_{ph} \lambda / h c + I_s) + 4 k T / (R_j \|(R_L + R_s))) \Delta f} \frac{1}{\Delta f} \quad (2.14)$$

Um das Verhältnis $\frac{S}{N}$ bei gegebener Photonenleistung P_{ph} zu maximieren, ist R_L zu vergrößern und die Bandbreite der Anwendung im Rahmen der Spezifikationen zu reduzieren.

2.2.3 Phototransistoren

Phototransistoren verfügen über eine interne Verstärkung des generierten Photostromes. Diese interne Verstärkung bewirkt bei gleicher eingestrahelter Intensität eine Erhöhung des Photostromes um den Betrag der Transistorverstärkung des vorwärts-aktiven⁶ Bereiches. Die Berechnung der Quanteneffizienz erfolgt gemäß Gleichung 2.1, so dass sich ein Wert für $\eta > 1$ ergeben kann; zum besseren Verständnis wird der Begriff der *effektiven* Quanteneffizienz eingeführt. Die effektiven Quanteneffizienzen des parasitären vertikalen pnp-Transistors des CMOS-Prozesses und des vertikalen npn-Transistors des BiCMOS-Prozesses wurden simuliert; die zur Device-Simulation benötigten Dotier-Profile der Transistoren wurden von AMS zur Verfügung gestellt. Abbildung 2.6 stellt die Dotierprofile und den schematischen Aufbau der beiden Transistoren dar.

⁶Basis-Kollektor-Spannung negativ, Basis-Emitter-Spannung positiv

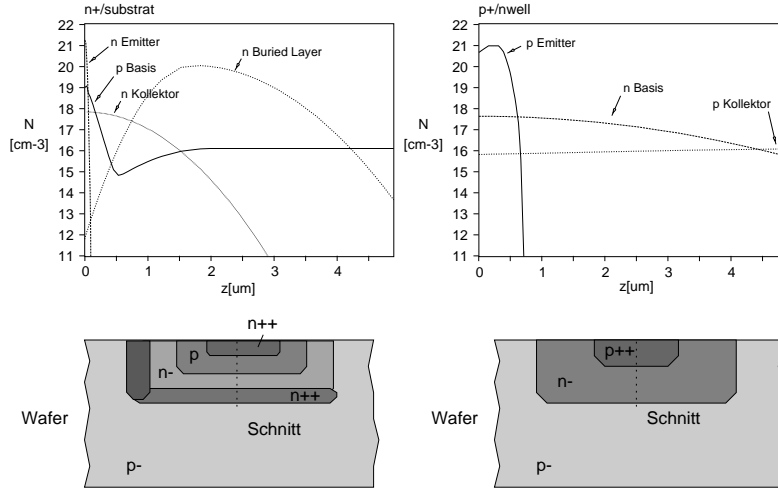


Abbildung 2.6: Dotierprofile der Phototransistoren

Die Simulation der spektralen Empfindlichkeit bestätigt die Verstärkung des generierten Photostromes, vergl. Abbildung 2.7. Der qualitative Verlauf der Effizienzen ist für die Lage der Kollektor-Basis-Sperrschicht des jeweiligen Transistors charakteristisch. Nach Abbildung 2.6 links liegt der pn-Übergang bei $z \approx 320 \text{ nm}$; das Maximum der spektralen Empfindlichkeit liegt bei $\lambda \approx 520 \text{ nm}$. Der pn-Übergang des parasitären vertikalen pnp-Transistors des CMOS-Prozesses befindet sich in einer Tiefe von $z \approx 4 \mu\text{m}$, vergl. Abbildung 2.6 rechts. Das Maximum der spektralen Empfindlichkeit liegt bei $\lambda \approx 640 \text{ nm}$. Der Verlauf der spektralen Empfindlichkeit ähnelt der p+/nwell-Diode und der nwell/sub-Diode.

Besonders hervorzuheben ist die Änderung der Verstärkung unter Änderung der Bestrahlungsintensität. Die Kurve *nnp* stellt die Höhe der effektiven Quanteneffizienzen des npn-Phototransistors für die Bestrahlungsintensitäten $J_{ph} = 1 \text{ W/m}^2$, $J_{ph} = 0.5 \text{ W/m}^2$ und $J_{ph} = 0.1 \text{ W/m}^2$ dar, die Kurve *pnp* die Höhe der effektiven Quanteneffizienz des pnp-Phototransistors für $J_{ph} = 1 \text{ W/m}^2$.

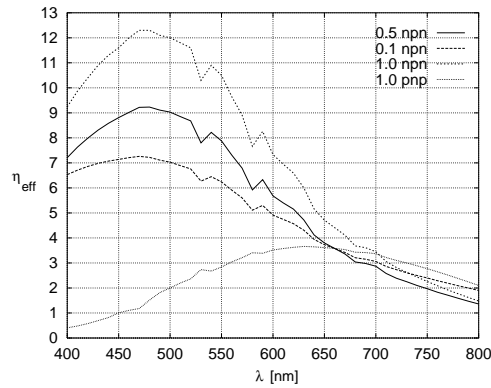


Abbildung 2.7: Quanteneffizienzen der Phototransistoren

Die Sprungantwortzeiten der Phototransistoren wurden ebenfalls in [Schneider 98] behandelt und sind in Tabelle 2.2 zusammengefasst. Die Höhe der jeweiligen Antwortzeit ist aufgrund der Abschätzung der Ladungsträger-Lebensdauern in der Simulation eher überbewertet. Angaben der Firma über den genauen Wert der Lebensdauern wurden nicht gemacht. Es ist jedoch

davon auszugehen, dass die tatsächlichen Sprungantwortzeit um bis zu einer Größenordnung geringer ist. Abbildung 2.8 zeigt das Ersatzschaltbild eines Phototransistors. An dieser Stelle

	nnp	pnp
τ [μs]	540	600

Tabelle 2.2: Sprungantwortzeit der Phototransistoren aus [Schneider 98]

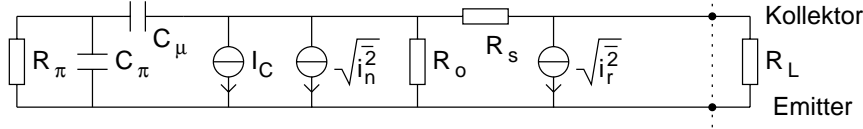


Abbildung 2.8: Ersatzschaltbild eines Phototransistors

seien nur die beiden Stromquellen I_C und $\sqrt{i_n^2}$ hervorgehoben, eine ausführliche Erläuterung des Kleinsignalersatzschaltbildes eines Bipolartransistors erfolgt zum Beispiel in [Grey 84].

Die Ladungstrennung der durch Absorption erzeugten Ladungsträger-Paare erfolgt hauptsächlich innerhalb der Verarmungszone der Basis-Kollektor-Sperrschicht. Zur Einhaltung der Ladungsträger-Bilanz werden zum Ausgleich der in die Basis gelangenden Majoritäten-Ladungsträger von dem Emitter Minoritäten-Ladungsträger in die Basis injiziert. Diese Injektion verursacht eine Basis-Emitter-Spannung, die den Transistor in den vorwärts-aktiven Bereich führt. Dementsprechend entsteht ein um die Verstärkung des Transistors erhöhter Kollektorstrom, der zusätzlich noch den Sperr-Sättigungsstrom der Basis-Kollektor-Sperrschicht enthält. Ein Phototransistor verstärkt daher den in der Basis-Kollektor-Sperrschicht entstehenden Photostrom um die Verstärkung h_{FE} :

$$I_C = (I_{ph} + I_{SC})(1 + h_{FE}) \quad (2.15)$$

$$\overline{i_n^2} = 2e((I_{ph} + I_{SC})(1 + h_{FE}) + I_C) \Delta f \quad (2.16)$$

h_{FE} ist die Gleichspannungs-Verstärkung, I_{SC} der Sättigungsstrom der Basis-Kollektor-Sperrschicht und I_{ph} der nach Gleichung 2.2 generierte Photostrom.

2.2.4 Gegenüberstellung der Photodetektoren

Die Eigenschaften der in den verwendeten Prozessen implementierbaren Photodetektoren werden in Tabelle 2.3 zusammengefasst. Im Verlauf der Entwicklung des Wellenfrontsensor-ASIC wurden drei verschiedene Photodetektoren implementiert, um das Verhalten zu evaluieren: die n+/sub-Photodiode, die p+/nwell-Photodiode und der npn-Phototransistor. Die Wahl der p+/nwell-Photodiode, trotz ihrer geringen spektralen Empfindlichkeit im nahen IR, ist durch Notwendigkeiten der Schaltungstopologie und dem Reduzieren des parasitären Übersprechens benachbarter Photodetektoren begründet. In Kapitel 3 werden diese Zusammenhänge erläutert.

Typ	Prozess	$\eta_e, \lambda = 632 \text{ nm}, J_{ph} = 1 \text{ W/m}^2$	$\tau \text{ [ns]}$
n+/sub	CMOS	≈ 0.6	1.5
p+/nwell	CMOS	≈ 0.25	0.62
nwell/sub	CMOS	≈ 0.65	3.7
pnp	CMOS	≈ 3	$540 \cdot 10^3$
npn	BiCMOS	≈ 5.2	$600 \cdot 10^3$

Tabelle 2.3: Vergleich der realisierbaren Photodetektoren

2.3 Der MOS-FET

Der MOS-FET ist das wichtigste Element der CMOS-Prozess-Technologie. Die Erforschung des Feldeffektes zur Steuerung von Strömen in Halbleitern führte zu einem Wechsel von den länger bekannten Bipolar-Transistoren zu Feldeffekt-Transistoren, die eine im Vergleich sehr viel geringere laterale Fläche zu ihrer Implementierung in der Silizium-Planar-Technologie benötigen. Die Hochintegration mikroelektronischer Schaltungen (VLSI⁷) wurde erst durch diese Prozess-Technologie ermöglicht. Auf der ohne Prozessfehler zu realisierenden Fläche konnte die zur Entwicklung komplexer digitaler Schaltungen erforderliche Anzahl von Transistoren angeordnet werden.

Ein weiterer Vorteil in der Verwendung von Feldeffekt-Transistoren entsteht durch die Möglichkeit zur leistungslosen Steuerung der Transistor-Ströme. Dadurch verringert sich die Verlustleistung beim Betrieb der Schaltung, so dass, wie z.B. bei älteren ECL⁸-Schaltungen, die problematische Wärmeableitung entfällt. Der große Unterschied der im Vergleich zur Bipolar-Technologie geringen Schaltfrequenzen der Feldeffekt-Transistoren ist im Verlauf der letzten Jahre durch die Reduzierung der minimalen Gate-Länge der FET-Strukturen nahezu egalisiert worden. Sogenannte *Deep Submicron*-Geometrien mit geringen Transit-Zeiten der Ladungsträger sind Gegenstand der Forschung.

Die CMOS-Technologie, in der komplementäre Feldeffekt-Transistoren zur Verfügung stehen, ist zuerst in digitalen Schaltungen eingesetzt worden, in denen die Qualität der MOS-FET-Charakteristika geringeren Ansprüchen unterliegt. Seit dem Einsatz der Feldeffekt-Technologie in analogen Schaltungen ist jedoch genau diese Qualität für die Funktionalität der Schaltung in Bezug auf das Erfüllen von Spezifikationen ausschlaggebend.

Im Folgenden werden die Eigenschaften des MOS-FET in Anwendungen von optoelektronischen Schaltungen untersucht und besonders die Verarbeitung der geringen Signalströme von Halbleiter-Photodetektoren betrachtet. Das Groß-Signal- und Kleinsignal-Ersatzschaltbild eines MOS-FET in Sättigungs- und linearem Bereich wird zum Beispiel in [Geiger 90] ausführlich behandelt, in diesem Kapitel erfolgt eine Beschreibung für den Subthreshold-Betrieb, da MOS-FETs in dieser Betriebsart oft in den elektrischen Schaltungen des Wellenfrontsensor-ASIC angewandt werden.

2.3.1 Der Subthreshold-Betrieb

Ein MOS-FET wird in vier Betriebsarten verwendet: Subthreshold-Bereich, linearer Bereich, Sättigungsbereich und Cutoff-Bereich. An dieser Stelle folgt eine Erläuterung der Arbeitsweise

⁷Very Large Scaled Integration

⁸Emitter Coupled Logic: Differenzen-Verstärkung durch Emitter-gekoppelte Bipolar-Transistoren

des MOS-FET im Subthreshold-Bereich, da in der Entwicklung des Wellenfrontsensor-ASIC speziell dieser Bereich verwendet wird, um eine Signalverarbeitung der geringen Photoströme zu implementieren. Eine Modell zur Berechnung des Drain-Stroms des Subthreshold-Bereiches wird definiert und die Parameter des Modells durch Messungen quantifiziert.

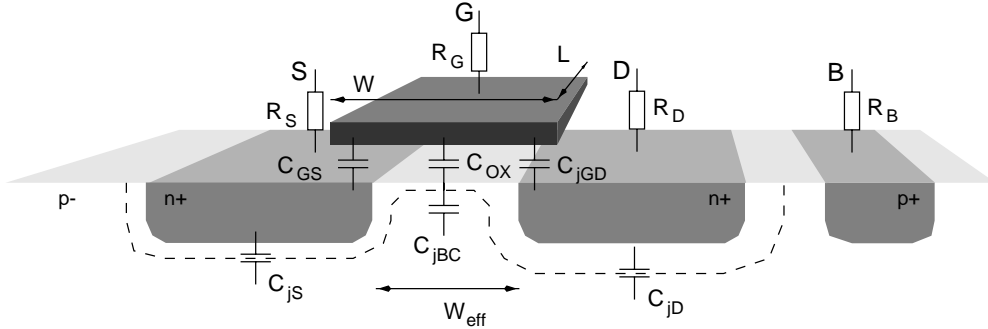


Abbildung 2.9: Schematische Ansicht eines NMOS-FET

Abbildung 2.9 zeigt eine schematische Ansicht eines MOS-FET. Die parasitären Kondensatoren und Widerstände des Bauelementes sind eingezeichnet, und die Lage der Raumladungszone angedeutet. Unter Anlegen einer Spannung V_{GS} zwischen Gate und Source entwickelt sich in dem Kanal zwischen Drain und Source eine Verarmungszone, in dem die freien Ladungsträger-Majoritäten in das Substrat gedrängt werden, und bei höheren Spannungen $V_{GS} > V_{TH}$ entsteht eine Inversion der Leitfähigkeit des Kanals durch Anreichern von Minoritäten. Im Gegensatz zum Stromtransport durch Drift dieser Minoritäten im Überschwellspannungs-Bereich unterliegt der MOS-FET im Subthreshold-Bereich $V_{GS} < V_{TH}$ dem Stromtransport-Mechanismus des Diffusionsstromes. Bei Verringerung der Gate-Source-Spannung V_{GS} unter die Threshold-Spannung V_{TH} fließt ein Drain-Strom, dessen Betrag nach [Vittoz 77] eine exponentielle Abhängigkeit von Drain-Bulk, Source-Bulk und Gate-Bulk-Spannung besitzt:

$$I_D = S I_{D0} e^{\frac{V_{GB}}{n V_T}} \cdot \left(e^{-\frac{V_{SB}}{V_T}} - e^{-\frac{V_{DB}}{V_T}} \right) \quad (2.17)$$

Es ist $S = \frac{W_{eff}}{L_{eff}}$ und bezeichnet das Verhältnis der effektiven Gate-Breite zur effektiven Gate-Länge; die "effektiven" Geometrien entstehen im Prozess durch Unterdiffusion der lateralen Maskierungs-Strukturen. I_{D0} und n sind prozessabhängige Parameter und bestimmen in der logarithmischen Darstellung die Steigung und den Ordinaten-Schnittpunkt des exponentiellen Verlaufs des Drain-Stromes im Subthreshold-Bereich. Diese beiden Parameter werden durch Messungen von MOS-FET-Strukturen des $0.8\mu\text{m}$ -CMOS-Prozesses bestimmt, in dem die Übertragungskennlinien $I_D = f(V_{GS})$ drei verschiedener MOS-FET-Geometrien gemessen und in dem relevanten Bereich $0.3 < V_{GS} < 0.7$ durch Einstellen der Parameter S , I_{D0} und n mit Gleichung 2.17 angenähert werden. Dabei werden unter Konstanz von S lediglich I_{D0} und n variiert, um eine Näherung des gemessenen Verlaufs zu erreichen, da S durch die Geometrie der MOS-FET-Strukturen selbst vorgegeben ist.

In Abbildung 2.10 ist der Verlauf der drei Messkurven von $M1$, $M2$ und $M3$ abgebildet. Zusätzlich sind die Annäherungen der Kennlinien durch $Sim1$, $Sim2$ und $Sim3$ dargestellt. Die Parameter der Näherungen sind in Tabelle 2.4 aufgelistet. S_n ist das nominale Verhältnis der Breite W und der Länge L der Maskenvorlage. Es zeigt sich eine gute Annäherung der beiden Übertragungs-Kennlinien von $M1$ und $M2$ für die Parameter $I_{D0} = 1.3 \text{ fA}$ und $n \approx 1.53$. Der Verlauf von $M3$ wird durch Wahl dieser beiden Parameter überbewertet, so daß eine Verringerung von S erforderlich ist. Dies ist möglicherweise ein Resultat der Ungenauigkeiten in der

Bestimmung der effektiven Weite W_L des Transistors. Der Parameter n ist nach [Vittoz 77] durch

$$n = 1 + \frac{1}{C_{ox}} \sqrt{\frac{e N_B \epsilon_{si}}{2(2\Phi_f - 5V_T + V_S)}} \quad (2.18)$$

gegeben, und unter Einsetzen der Werte⁹ $N_B = 74 \cdot 10^{15}/cm^{-3}$, $C_{ox} = 5.533 \cdot 10^{-8} F/cm^2$, $\Phi_f = V_T \ln N_B/n_i = 0.39 V$, $V_T = 25.2 mV$ und $n_i = 1.45 \cdot 10^{10}/cm^{-3}$ ergibt sich ein Betrag von $n \approx 1.51$; der theoretische Wert stimmt auf 1% mit dem angenäherten Wert überein. Die

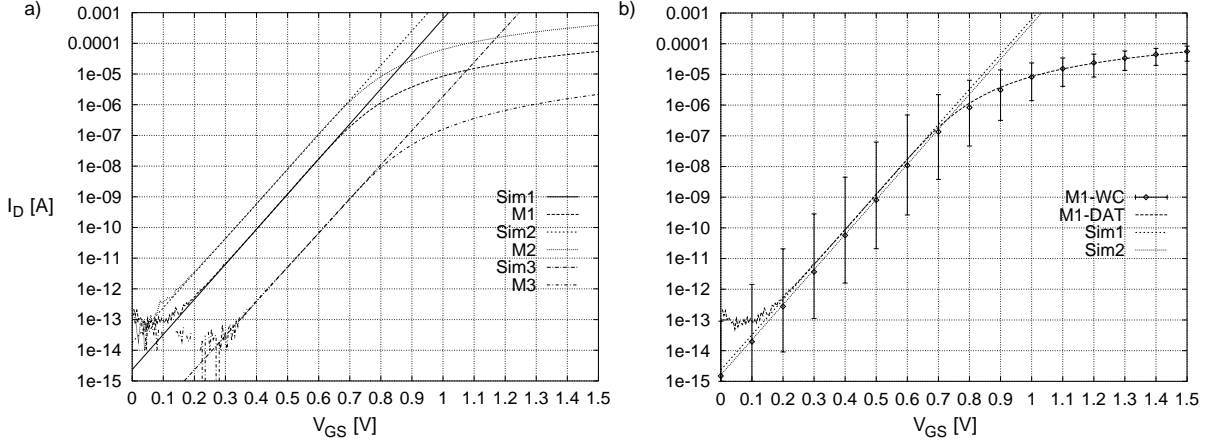


Abbildung 2.10: Der Drain-Strom eines MOS-FET im Subthresholdbereich

MOS-FET	W [μm]	W_{eff} [μm]	L [μm]	L_{eff} [μm]	S_n	S	I_{D0} [A]	n	λ [V^{-1}]
M1	2.0	1.2	0.8	0.66	1.81	1.81	$1.3 \cdot 10^{-15}$	1.52	0.30
M2	10.0	9.2	0.8	0.66	13.9	13.9	$1.3 \cdot 10^{-15}$	1.54	0.23
M3	2.0	1.2	10	9.86	0.12	0.0105	$1.3 \cdot 10^{-15}$	1.56	0.02

Tabelle 2.4: Näherungs-Parameter S , I_{D0} , n und λ

Übertragungskennlinie eines MOS-FET des $0.8\mu m$ -CMOS-Prozess kann im Subthreshold-Bereich für $0.3 < V_{GS} < 0.7$ durch Gleichung 2.17 mit den Parametern $I_{D0} = 1.3 fA$, $n \approx 1.53$ und $S = W_{eff}/L_{eff}$ hinreichend genau modelliert werden, solange $S > 1$ ist. Bei $S < 1$ ist eine Korrektur des Wertes S vorzunehmen.

In Teilbild b) der Abbildung 2.10 wird die Streuung der Prozess-Parameter in der $0.8\mu m$ -CMOS-Prozess-Technologie verdeutlicht. Die Kurve *M1-WC* entspricht den Ergebnissen der Simulationen eines MOS-FET durch *SpectreS*, unter Auswahl des *BSim3v2*-Modelles. Gleichzeitig ist die Messkurve aus Teilbild a) (*M1*) wiederholt. Die Fehlerbalken quantifizieren die Streuung der simulierten Drain-Ströme unter Auswahl der Eck-Parameter¹⁰ und verdeutlichen die mögliche Schwankung des Drain-Stromes im Subthreshold-Bereich um ca. 3 Dekaden. Am

⁹zum Teil aus [AMS 0.8]

¹⁰Die Eck-Parameter repräsentieren die im Prozess maximal auftretenden Streuungen der Prozess-Parameter. Dazu werden die Parameter-Sätze *tm*, *wz* und *wo* ausgewählt, entsprechend den typischen Parametern (*tm*), den Parametern, die einen "schwachen" NMOS-FET (*wz*) beschreiben, und den Parametern, die einen schwachen PMOS-FET beschreiben (*wo*).

Übergang in den Strong-Inversion-Bereich fällt diese Fehlerbreite entsprechend der nurmehr quadratischen Abhängigkeit des Drain-Stromes von der Gate-Source-Spannung auf unter eine Dekade ab. Zusätzlich sind zwei Annäherungen durch Gleichung 2.17 mit den Werten $S = 1.81$, $n = 1.52$ eingezeichnet, wobei der Strom I_{D0} für *Sim1* 1.3 fA und für *Sim2* 0.9 fA beträgt. Durch *Sim1* ist die Messkurve angenähert worden, durch *Sim2* die Simulations-Kurve der “typischen” Prozess-Parameter, die durch die Punkte der Fehlerbalken dargestellt werden.

Um den Effekt der Kanallängen-Modulation durch Variation der Spannung V_{DS} zu berücksichtigen, durch den aufgrund der Erweiterung der Raumladungszone an den Drain-Gebiet eine Verkürzung des Kanals entsteht, wird Gleichung 2.17 um einen entsprechenden Term erweitert.

$$I_D = S I_{D0} e^{\frac{V_{GS}}{n V_T}} \cdot \left(e^{-\frac{V_{SB}}{V_T}} - e^{-\frac{V_{DB}}{V_T}} \right) \cdot (1 + \lambda V_{DS}) \quad (2.19)$$

Der Parameter λ wird ebenfalls durch die Verarbeitung gemessener Daten bestimmt. Dazu werden die Ausgangskennlinien $I_D = f(V_{DS})$ der drei verschiedenen MOS-FET-Strukturen eines weiteren Test-Chips bei $V_{GS} = 0.4 \text{ V}$ gemessen und im relevanten Bereich $0.3 < V_{DS} < 5$ durch die Gleichung 2.19 angenähert, wie in Abbildung 2.11 dargestellt. Es ergeben sich die in Tabelle 2.4 referenzierten Werte für λ . $\lambda(\text{M3})$ sei an dieser Stelle hervorgehoben, da der Faktor der Kanallängen-Modulation sehr gering ist, eine Tatsache, die zum Entwurf hochkonstanter Stromquellen von Wichtigkeit ist. Dies ist durch die geringe Steigung der Kurven *M3* und *Sim3* zu erkennen. Im Entwurf des Wellenfrontsensor-ASIC werden als Stromquelle eingesetzte MOS-FET-Strukturen im Allgemeinen mit $S < 1$ implementiert.

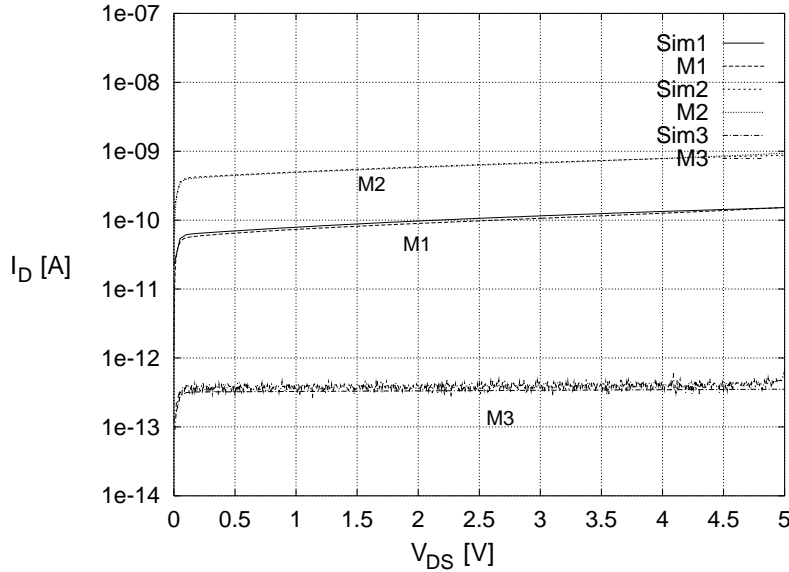


Abbildung 2.11: Annäherung der Kanalweiten-Modulation im Subthresholdbereich

Zusammenfassend lässt sich das Kleinsignal-Ersatzschaltbild des MOS-FET für den Betrieb im Subthreshold-Bereich aufstellen; es gilt $V_{BS} = 0 \text{ V}$.

Dabei ist für z.B. M1

$$\begin{aligned} I_D &= 1.81 \cdot 1.3 \text{ fA} e^{\frac{V_{GS}}{1.52 V_T}} \cdot \left(1 - e^{-\frac{V_{DS}}{V_T}} \right) \cdot (1 + 0.3/V V_{DS}), \\ g_m &= \frac{dI_D}{dV_{GS}} = \frac{1}{n V_T} I_D, \end{aligned} \quad (2.20)$$

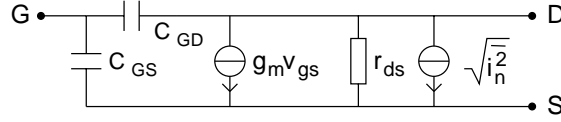


Abbildung 2.12: Das Kleinsignal-Ersatzschaltbild eines MOS-FET nach [Allen 87b]

$$r_{ds} = \left(\frac{dI_D}{dV_{DS}} \right)^{-1} \approx (\lambda I_D)^{-1}, \quad (2.21)$$

$$\overline{i_n^2} = \left(\frac{8}{3} kT g_m + \frac{KF I_D}{C_{ox} L^2} \frac{1}{f} \right) \Delta f. \quad (2.22)$$

Die Beziehung 2.22 gibt dabei den Anteil des Rauschens wieder. In einem Feldeffekt-Transistor existieren zwei Ursachen des Rauschens: das sog $1/f$ -Rauschen und das thermische Rauschen. Das thermische Rauschen beschreibt den Rauschanteil, der durch thermische Bewegung der Ladungsträger entsteht und jedem Widerstand, also auch dem Kanal des FET zu eigen ist. Das $1/f$ -Rauschen ist eine spezielle Eigenschaft der Feldeffekt-Transistoren und entsteht durch das statistische niederfrequente Besetzen und Verlassen von Ladungsträger-Energiezuständen in der Übergangsschicht zwischen Si und SiO₂ [Gregorian 86]. Dieser Anteil des Rauschens eines MOS-FET wird maßgeblich durch den Prozessparameter KF bestimmt.

Eine weitere zu untersuchende Eigenschaft des MOS-FET ist das sogenannte *Mismatching*, das die Unterschiedlichkeit der elektrischen Eigenschaften identischer Strukturen auf einem Wafer beschreibt. Während die vorangehend beschriebenen Eck-Parameter die Streuung der charakteristischen Eigenschaften zwischen einzelnen Batches¹¹, also die globale Streuung der Prozessparameter, beschreibt, ist das Mismatching auf die Übereinstimmung *benachbarter* Strukturen bezogen. Änderungen in der Gate-Spannung eines MOS-FET in Bereichen von einigen Millivolt bewirken im Subthreshold-Betrieb aufgrund der exponentiellen Abhängigkeit des Drain-Stromes von der Gate-Spannung entsprechend große Unterschiede:

$$\frac{I_{D1}}{I_{D2}} = e^{\frac{\Delta V_{GS}}{n V_t}} \quad (2.23)$$

Die Entwicklung der integrierten Schaltung erfolgt unter Layout-Techniken, die das Mismatching reduzieren. Ein geringer Unterschied ist jedoch nicht zu vermeiden, so dass die Funktionalität der Schaltung selbst Auswirkungen des Mismatchings kompensieren können muss. Die quantitative Aussage über das Mismatching der MOS-FET-Strukturen des Wellenfrontsensor-ASIC erfolgt in Kapitel 6.

Die in diesem Kapitel beschriebenen Elemente der CMOS- und BiCMOS-Technologie wurden zur Entwicklung des Wellenfrontsensor-ASIC eingesetzt. Die Berechnung der Funktion der verschiedenen Schaltungen des ASIC erfolgte im Allgemeinen unter Verwendung von Gleichung 2.19, da die MOS-FET des ASIC aufgrund der geringen Werte der in der Anwendung als Wellenfront-Sensor auftretenden Photoströme im Subthreshold-Bereich betrieben werden. Im folgenden Kapitel wird die Entwicklung des ASIC erläutert.

¹¹Ein Batch entspricht einem Verbund von bis zu 12 Wafern, die den Herstellungsprozess gleichzeitig durchlaufen

Kapitel 3

Die Entwicklung des ASIC

Im folgenden Kapitel wird die Entwicklung der elektronischen Schaltungen und des Layouts des Wellenfrontsensor-ASIC vorgestellt. Ausgehend von den in Kapitel 1.6 genannten Spezifikationen werden Konzepte vorgestellt, die im Rahmen des Entwurfs zweier Prototypen verbessert und als Grundlage der Systemarchitektur des ASIC eingesetzt wurden.

Das Kapitel ist dazu in vier Abschnitte gegliedert. Im ersten Abschnitt erfolgt eine Übersicht über das System. Die detaillierte Erläuterung der Entwicklung des Wellenfrontsensor-ASIC-Prototypen, im Folgenden als *HSSX*¹ bezeichnet, erfolgt unter Betrachtung des Schaltungsentwurfs, der Layout-Gestaltung und der Systemintegration im zweiten Abschnitt. Im dritten und vierten Abschnitt werden die beiden Teststudien *Prototyp 1* und *Prototyp 2* des Projekts in einer kurzen Übersicht vorgestellt und deren Implementierungen des Algorithmus zur Positionsdetektion der Fokalfunkte erläutert.

3.1 Systemübersicht

Nach den Ausführungen des Kapitels 1.6 besteht die Funktion des Wellenfrontsensor-ASIC darin, die lateralen Verschiebungen von Fokalfunkten zu detektieren, diese in Koordinaten für X- und Y-Richtung zu diskretisieren und die Positionsinformationen an das System zu übergeben. Die Fokalfunkte, in einer Matrix angeordnet, erzeugen je einen charakteristischen Photonenfluss entsprechend Gleichung 1.13. Im Allgemeinen existiert genau eine Position x^* und y^* , an der maximale Intensität vorliegt. Der ASIC detektiert die Position und diskretisiert den analogen Wert in Abschnitte. Die zu realisierende Breite dieser Abschnitte Δx und Δy hängt mit der Genauigkeit in der Rekonstruktion der Wellenfront zusammen, siehe Kapitel 1.6.1, und beträgt $17 \mu\text{m}$. Die diskretisierten Daten werden an ein externes System übergeben, in dem der Least-Square-Fit zur Berechnung der Zernike-Koeffizienten durchgeführt wird.

Der Wellenfrontsensor-ASIC ist in sog. *Cluster* unterteilt; jedes Cluster detektiert die laterale Verschiebung eines Fokalfunktes. In Abbildung 3.1 ist der schematische Aufbau der drei Prototypen dargestellt und die Anordnung der Cluster in der *Cluster-Matrix* verdeutlicht. Jedes Cluster

¹Hartmann-Shack-Sensor-eXtreme

erzeugt X- und Y-Daten zur Beschreibung der diskreten Fokuspunkt-Position und übergibt diese in einer Daisy-Chain-Architektur² an das nächste Cluster der Spalte oder Zeile. Die Daten wer-

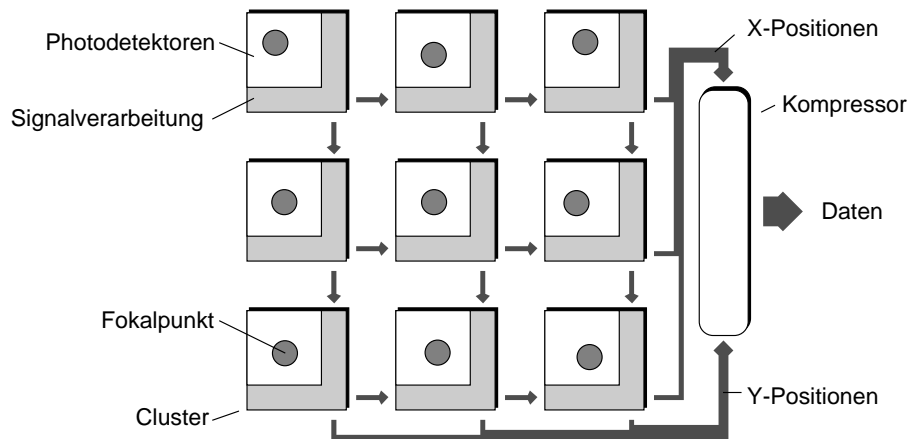


Abbildung 3.1: Die Architektur des Wellenfrontsensor-ASIC

den über eine auf dem Prototypen befindliche Schnittstelle an das externe Datenverarbeitungssystem übertragen, in dem die Berechnung des Least-Square-Fit vorgenommen wird. Die Beschreibung des externen Systems ist Bestandteil des Kapitels 4. Die anfallende Datenmenge der Cluster-Matrix wird im *HSSX* durch einen *Kompressor* komprimiert. Diese Einheit ist nur in dem *HSSX* implementiert, da *Prototyp 1* und *Prototyp 2* aufgrund der geringeren Anzahl der Cluster³ eine geringere Datenmenge erzeugen, die ohne Kompression übertragen werden kann.

3.2 Der HSSX

Abbildung 3.2 verdeutlicht den Aufbau eines der 256 Cluster des *HSSX*. In der *Photodetektor-*

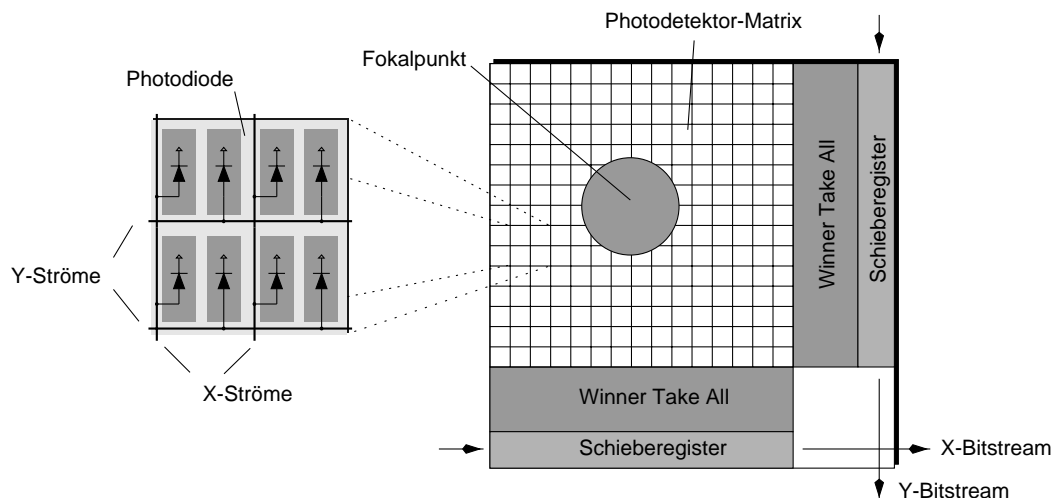


Abbildung 3.2: Der Aufbau eines Cluster

²Kaskadierte Schieberegister, analog zu dem angelsächsischen Begriff einer "Gänseblümchen-Kette"

³Der *HSSX* verfügt über 16×16 Cluster, Prototyp 1 über 4×4 , Prototyp 2 über 3×3

Matrix sind in einer äquidistanten Anordnung Pixel realisiert, die aus jeweils zwei p+/nwell-Photodioden bestehen. Die rechten Photodioden der Pixel einer Matrix-Zeile sind miteinander verbunden, sowie die linken Photodioden der Pixel einer Matrix-Spalte. Dadurch ist es möglich, Photodetektor-Streifen zu realisieren, die den Photostrom, der bei Auftreffen eines Fokalkpunktes auf die Matrix entsteht, in Zeilen- und Spalten-Anteile diskriminieren. An der rechten und unteren Seite des Cluster befinden sich Schaltungsmodule, die durch analoge Signalverarbeitung den höchsten Spalten- und Zeilenstrom detektieren können. Das *Winner-Take-All*-Modul (WTA-Modul) erzeugt dabei digitale Ausgangssignale, die in das *Schieberegister* geladen werden können. In einem sich der Detektion anschließenden Auslesevorgang werden die Schieberegister aller Cluster der 16 Zeilen und 16 Spalten der Cluster-Matrix miteinander verbunden und die Signale der Winner-Take-All-Module ausgelesen. Nach Abbildung 3.1 komprimiert der *Kompressor* die anfallenden Daten und stellt sie dem externen System zur Verfügung.

Die Funktionalität des Winner-Take-All-Moduls sei durch Abbildung 3.3 verdeutlicht. In der Abbildung sind zur Übersichtlichkeit nur 64 Pixel des Cluster (entsprechend 8×8 Zeilen und Spalten) dargestellt⁴. Der Fokalkpunkt repräsentiert eine Intensitätsverteilung J_{ph} , die durch

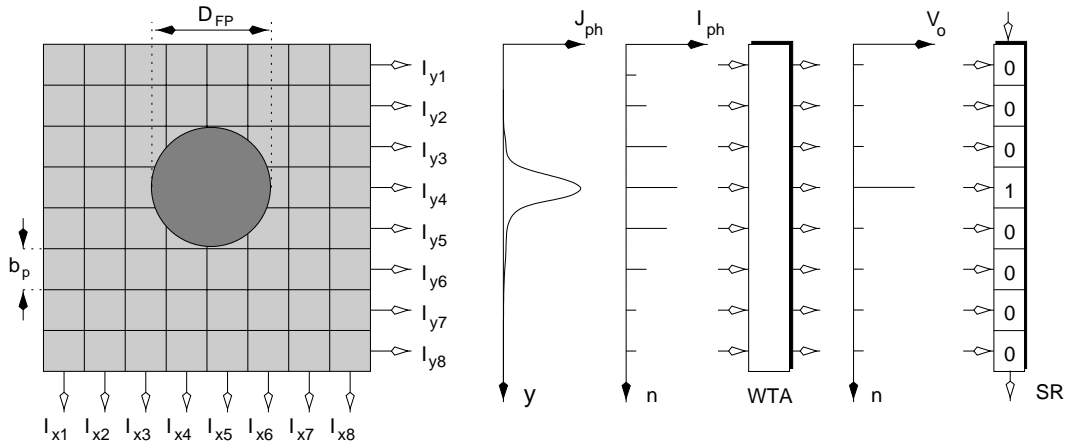


Abbildung 3.3: Die Diskretisierung der Photoströme in digitale Spannungen

die Airy-Funktion beschrieben werden kann, vergleiche Gleichung 1.12. Der Durchmesser D_{FP} des Fokalkpunktes beträgt nach Formel 1.8 ca. $126 \mu m$, daher werden mehrere Pixel gleichzeitig bestrahlt. Die Kantenlänge eines Pixels b_p beträgt im *HSSX* $17.05 \mu m$, so dass ca. 7 Pixel innerhalb der Grenzen des Fokalkpunktes einen Photostrom generieren. Die Beträge der Zeilenströme I_y sind als Funktion $I_{ph}(n)$ dargestellt. Unter Annahme einer idealen Fokalkpunkt-Gestalt verhalten sich die Anteile der betrachteten 7 Zeilen-Ströme nach

$$I_{y(4+z)} = I_{y4} \frac{1}{3.5 b^2} \int_{-3.5 b_p}^{3.5 b_p} \int_{-b_p/4+z \cdot b}^{b_p/4+z \cdot b} Z(x, y) dy dx \quad z \in \{\pm 1, \pm 2, \pm 3\} \quad (3.1)$$

mit $Z(x, y)$ gemäß Gleichung 1.14 wie $I_{y4} = 100\%$, $I_{y5} = I_{y3} = 74\%$, $I_{y6} = I_{y2} = 27\%$ und $I_{y7} = I_{y1} = 3.3\%$. Das WTA-Modul erkennt innerhalb der spezifizierten Zeitdauer von ca. 1 ms den höchsten Strombetrag in I_{y4} und erzeugt ein entsprechendes Bitmuster, das in das Schieberegister (SR) geladen wird. Die absoluten Beträge der Photoströme werden nach den Spezifikationen aus Kapitel 1.6 bei Fokalkpunkt-Photonenleistungen von ca. 200 pW in dem Bereich

⁴Die Cluster des *HSSX* verfügen über 19×19 Zeilen und Spalten mit 361 Pixeln, entsprechend 722 Photodetektoren

von einigen pA liegen. Um das Ausbleiben einer Photonenbestrahlung von definierter Intensität detektieren zu können, wird das WTA-Modul um eine Zelle erweitert, über die ein einstellbarer Strom mit den Photoströmen der Detektor-Matrix verglichen werden kann. Die Zelle wird als *Dummy*-Zelle bezeichnet. Sie ist in Abbildung 3.3 nicht eingezeichnet.

3.2.1 Das Timing des HSSX

Die Systemspezifikation erfordert die Wiederholrate von 1 kHz, so dass sich die Detektions- und Auslesedauer über den Zeitraum von 1 ms erstreckt. Das System verwendet davon eine Dauer von ca. 980 μs zur Detektion des stärksten Eingangs-Stromes durch die WTA-Module und den Zeitraum von 20 μs zum Laden der WTA-Informationen in die Schieberegister und zum anschließenden Übertragen der Daten an das externe System. Das Schieben der Daten erfordert zum vollständigen Übertragen aller Informationen der 16 Cluster einer Zeile bei jeweils 20 WTA-Zellen⁵ 320 Schiebeimpulse, so dass sich bei einer Ladedauer von ca. 10 μs in den verbleibenden 10 μs des Auslesevorganges ein Systemtakt von 32 MHz ergibt. Die Daten der jeweils 16 Cluster einer Zeile und Spalte werden während des Schiebens in einer Pipeline⁶-Architektur parallel komprimiert und an die Ausgangs-Ports des *HSSX* übertragen. Der Auslesevorgang (Laden und Schieben/Komprimieren) kann nicht zeitgleich zu der Detektion erfolgen, da durch das Übersprechen⁷ der digitalen Schaltungen über die Versorgungsleitungen und das Substrat die analogen WTA-Schaltungen aufgrund der hohen Datenübertragungs-Rate empfindlich gestört würden. Im Folgenden werden die Pixel, das WTA-Modul, die Schieberegister und der Kompressor unter Betrachtung der zeitlichen Restriktionen beschrieben.

3.2.2 Das Pixel

Ein Pixel des *HSSX* umfasst zwei p+/nwell-Photodioden auf einer Fläche von $17 \times 17 \mu\text{m}^2$. Nach den Ausführungen des Kapitel 2.2.2 ist die spektrale Empfindlichkeit dieser Photodiode geringer als die der anderen beiden Dioden-Typen des CMOS-Prozesses. Die Wahl der p+/nwell-Photodiode ist zum einen motiviert durch das geringere Übersprechverhalten benachbarter Strukturen, da durch den zweiten etwas tiefer im Substrat liegenden pn-Übergang zwischen N-Wanne und Substrat die statistisch diffundierenden Ladungsträger eliminiert werden. Zum anderen erfordert die in Kapitel 3.2.3 vorgestellte Schaltungstopologie des WTA-Moduls eine Photodiode als Stromquelle, die durch die Kathode mit dem höchsten Potential der Schaltung verbunden ist. Die n+/sub- und die nwell/sub-Photodiode verwenden das Substrat des Wafers als Anode; das Substrat ist per Definition mit Masse verbunden (vergleiche Abbildung 2.1).

3.2.3 Das Winner-Take-All Modul

Das WTA-Modul stellt die zentrale Funktionalität des Wellenfrontsensor-ASIC dar. Die Detektion der lateralen Position des Fokalkpunktes eines Cluster geschieht durch die Diskriminierung des stärksten unter den verschiedenen durch den Fokalkpunkt generierten Photoströmen. Diese Problematik ist transferierbar in das Problem, in einem Vektor mit den Elementen $\{s_1, s_2, s_3, \dots, s_n\}$ das Element s_i mit $s_i > s_n, \forall n \neq i$ zu finden. Strategien zur Lösung des Problems

⁵19 Photodetektor-Zeilen oder -Spalten und jeweils eine *Dummy*-Zelle

⁶Zeitgleiches Verarbeiten von repetitiven Daten durch aufeinanderfolgende Funktionen

⁷Als Übersprechen werden alle Formen der unerwünschten gegenseitigen Einflussnahme der verschiedenen Schaltungsmodule einer Schaltung bezeichnet

sind in der digitalen Signalverarbeitung etabliert und es bestehen verschiedenste Ansätze der Algorithmik.

In diesem Projekt wird der Ansatz der analogen Signalverarbeitung gewählt. Durch die Anwendung der optoelektronischen Eigenschaften von Elementen der integrierten Schaltungstechnik ist es möglich, eine instantane Parallelisierung zu etablieren, die eine besonders schnelle Verarbeitung von Signalen ermöglicht. Als einfache Beispiele seien dabei die arithmetischen Funktionen wie Addition, Multiplikation, Division oder Logarithmierung genannt, wie auch komplexe Algorithmen wie Normalisierung, Mittelwerts-Bildung, Vektor-Längen-Berechnung, o.ä. [Vittoz 96]. Der prinzipielle Ansatz dieser Algorithmen beruht auf der Anwendung Kirchhoffscher Regeln⁸ und linearer oder nichtlinearer Kennlinien von passiven und aktiven Elementen. Als neue Klasse der analogen Signalverarbeitung ist die Entwicklung *neuronaler Netze* Gegenstand der Forschung, in der mehrdimensionale Netze aus Neuronen durch mehr oder weniger vollständige Verschaltungen untereinander einen Algorithmus ausführen; eine Technik, die in der digitalen Signalverarbeitung seit längerem etabliert ist, jedoch in der Analog-Technik erst seit Qualifizierung der analogen VLSI Einzug hält [Langeheine 98]. Als Vorteil der analogen Signalverarbeitung sei die hohe Signalverarbeitungs-Geschwindigkeit angeführt, als Nachteil die Ungenauigkeit der Lösungen durch inhärente Einflüsse der analogen Signalverarbeitung wie Mismatching⁹, Rauschen und Übersprechen. Im Folgenden wird der analoge Algorithmus des im Wellenfront-sensor-ASIC angewandten WTA-Moduls vorgestellt. Die Basis der Schaltung wird betrachtet und Modifikationen der Basis-Funktionalität zur Einhaltung der in Kapitel 1.6 definierten Zeit-Spezifikationen werden erörtert.

Das Grundprinzip des WTA-Moduls

Das WTA-Modul repräsentiert eine Kaskade von Zellen, die durch einen Signalpfad miteinander verbunden sind. Jede Zelle n erhält ein Eingangssignal I_{in} und erzeugt ein Ausgangssignal I_{on} . Die Kommunikation mit den übrigen Zellen findet durch das Übertragen eines Signals über einen einzelnen Signalpfad statt. Jede Zelle enthält zwei MOS-FET und ist daher in der Basis-Version mit geringem Platzbedarf zu implementieren. Abbildung 3.4 zeigt die Verschaltung und den Aufbau der MOS-FET von 3 WTA-Zellen.

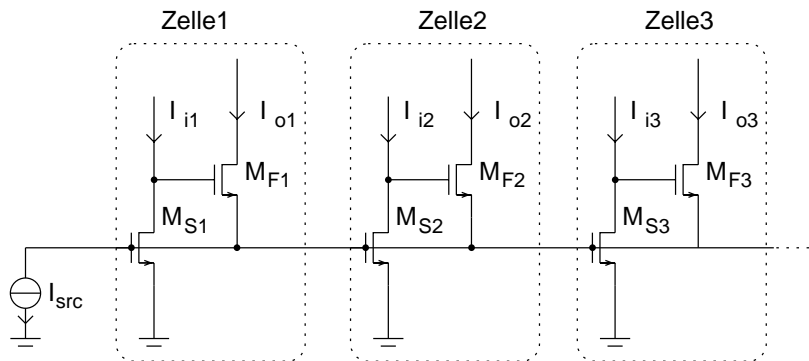


Abbildung 3.4: Das Grundprinzip des Winner-Take-All-Moduls

Eine jede Zelle verwendet einen *Sens*-Transistor M_{Sn} und einen *Flow*-Transistor M_{Fn} . M_{Sn} erfährt als Drain-Strom den Eingangsstrom I_{in} , testet also das Eingangssignal. M_{Fn} wird durch

⁸Summe aller umlaufenden Spannungen einer Masche = 0, Summe aller Knotenströme = 0

⁹Mismatching bezeichnet das unterschiedliche Verhalten identischer Strukturen oder Elemente der Signalverarbeitung

das Drain-Potential des M_{Sn} gesteuert und kontrolliert den Stromfluss von I_{src} , der als Ausgangssignal I_{on} verwendet wird. Durch die Gate-Verbindung aller M_{Sn} ist gewährleistet, dass die Transistoren die gleiche Ausgangscharakteristik besitzen. Unter Vernachlässigung des Body¹⁰-Effekts der M_{Fn} und nach

$$I_D = S I_{D0} e^{\frac{V_{GS}}{nV_T}} \cdot \left(1 - e^{-\frac{V_{DS}}{V_T}}\right) \cdot (1 + \lambda V_{DS}) \quad (3.2)$$

befindet sich der Arbeitspunkt aller M_{Sn} auf einer Ausgangskennlinie $I_D(V_{DS})$. Durch die Höhe des Drain-Stromes $I_{D(MSn)}$ wird zur Einhaltung des Gleichgewichts die Drain-Source-Spannung V_{DS} jedes M_{Sn} vorgegeben. Die Drain-Source-Spannung steuert die Gate-Source-Spannung des M_{Fn} . Nach Gleichung 3.2 bestimmt die Gate-Source-Spannung von M_{Fn} die Höhe des Drain-Stromes $I_{D(MFn)}$ und damit I_{on} . Aufgrund der flachen Steigung von I_D als Funktion von V_{DS} erzeugt ein geringer Unterschied der I_{in} einen großen Unterschied der I_{on} . Dieser Unterschied wird durch die erzwungene Differenzbildung der Ströme I_{MFn} durch $\sum_n I_{D(MFn)} = I_{src}$ verdeutlicht. Das Grundprinzip des WTA-Moduls basiert auf der Steigung der Ausgangskennlinie

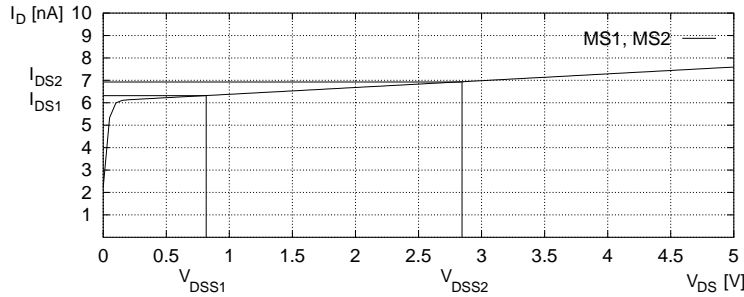


Abbildung 3.5: Ein geringer Unterschied von $I_{D(MS1)}$ und $I_{D(MS2)}$ erzeugt einen großen Unterschied von $V_{DS(MS1)}$ und $V_{DS(MS2)}$

der Sens-MOS-FET. Nach Gleichung 3.2 wird die charakteristische Eigenschaft der Kanallängen-Modulation eines MOS-FET durch den Parameter λ bestimmt und kann durch die Wahl eines im Vergleich zur Breite sehr langen Kanals ($S < 1$) eingestellt werden. Nach den Ausführungen von Kapitel 2.3.1 beträgt der Wert λ für einen MOS-FET mit $W/L = 2/10$ ca. 0.03. Die Änderung der Drain-Source-Spannung mit dem Drain-Strom berechnet sich aus

$$\frac{\Delta I_D}{\Delta V_{DS}} = \frac{\delta I_D}{\delta V_{DS}} \approx \lambda I_D \quad \lambda \ll 1 \quad (3.3)$$

zu

$$\Delta V_{DS} = \frac{\Delta I_D}{I_D} \frac{1}{\lambda}, \quad (3.4)$$

d.h. die Verstärkung einer Drain-Stromänderung $\Delta I_D/I_D$ beträgt λ^{-1} . Die Qualität des WTA-Moduls in der Unterscheidung der Eingangsströme I_{in} in Bezug auf die Diskriminierungsgenauigkeit und das Transientenverhalten ist ausschlaggebend für die Funktionalität des Wellenfrontsensor-ASIC. Im Folgenden wird das Transientenverhalten und die Diskriminierungsgenauigkeit des WTA-Moduls anhand zweier WTA-Zellen quantifiziert.

Abbildung 3.6 a) stellt eine WTA-Zelle mit Stromquelle I_{src} dar, deren Eingangsstrom momentan der stärkste der I_{in} ist. Der Arbeitspunkt der Schaltung, also die Größe der Ruhestrome der MOS-FET unter Anliegen von statischen Eingangssignalen, wird durch die Berechnung von $I_{D(MF1)}$ und $I_{D(MS1)}$ festgestellt.

¹⁰Veränderung der Schwellspannung eines MOS-FET durch Differenz von Source- und Substrat-Potential

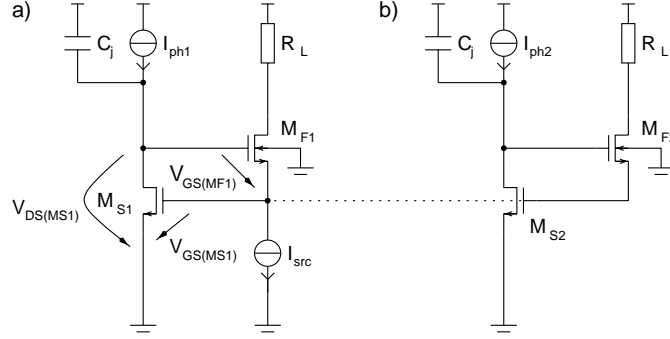


Abbildung 3.6: Arbeitspunkt-Berechnung des WTA-Moduls

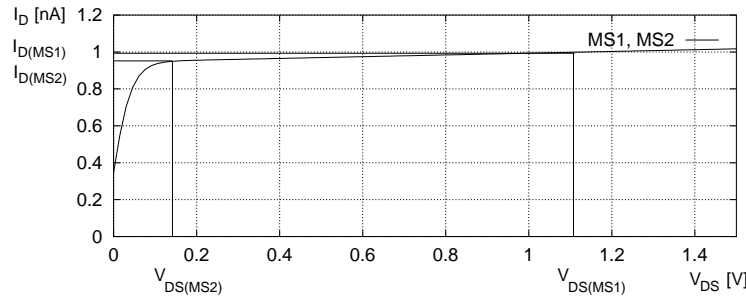
Aus der Beziehung $V_{GS(MF1)} + V_{GS(MF2)} = V_{DS(MS1)}$ ergibt sich unter Gleichsetzung von $I_{D(MS1)} = I_{ph1}$ und $I_{D(MF1)} = I_{src}$ und Vernachlässigung des Body-Effektes von M_{F1} für M_{S1} eine Gate-Source- und eine Drain-Source-Spannung von

$$\begin{aligned} V_{GS(MS1)} &= \ln \left(\frac{I_{ph1}}{I_0} \cdot \frac{1}{S(1 - e^{-V_{DS(MS1)}/V_T})(1 + \lambda V_{DS(MS1)})} \right) n V_t \quad \text{und} \\ V_{DS(MS1)} &= V_{GS(MS1)} + \ln \left(\frac{I_{src}}{I_0} \cdot \frac{1}{S(1 - e^{-(V_{dd} - R_L I_{src} - V_{GS(MS1)})/V_T})(1 + \lambda (V_{dd} - R_L I_{src} - V_{GS(MS1)}))} \right) n V_t \end{aligned} \quad (3.5)$$

Dieses Gleichungssystem lässt sich numerisch lösen, und unter Wahl der Parameter $I_0 = 1 \text{ fA}$, $\lambda = 0.05 \text{ V}^{-1}$, $I_{ph1} = 1 \text{ nA}$, $I_{ph2} = 0.95 \text{ nA}$, $I_{src} = 100 \text{ nA}$, $n = 1.5$, $S = 2$ und $V_t = 25 \text{ mV}$ ergeben sich Werte von $V_{GS(MS1)} = 542 \text{ mV}$ und $V_{DS(MS1)} = 1.162 \text{ V}$. Der Arbeitspunkt der “verlierenden” WTA-Zelle in Teilbild b) berechnet sich unter der Annahme, dass $I_{D(MF2)}$ sehr viel geringer als $I_{D(MF1)}$ ist. Diese Annahme ist legitimiert, da M_{F2} als Gate-Source-Spannung einen negativen Spannungswert erfährt, der sich aufgrund des geringeren Drain-Potentials von M_{S2} ergibt: M_{F2} befindet sich im Cutoff-Bereich. $V_{DS(MS2)}$ beträgt

$$V_{DS(MS2)} = \left(\frac{I_{ph2}}{I_0 S e^{V_{GS(MS1)}/n V_t} (1 - e^{-V_{DS(MS2)}/V_t}} - 1 \right) \lambda^{-1} \quad (3.6)$$

Diese Gleichung lässt sich ebenfalls numerisch lösen und für die zuvor berechneten Werte ergibt sich $V_{DS(MS2)} = 0.152 \text{ V}$. Die beiden errechneten Drainströme von M_{S1} und M_{S2} sind in Abbildung 3.7 dargestellt.

Abbildung 3.7: Die Kennlinie von M_{S1} und M_{S2}

Die Schaltung befindet sich in einem stabilen Ruhezustand, bis eine Änderung der Eingangssignale eintritt. Bei Zutreffen der Bedingung $I_{ph1} < I_{ph2}$ entsteht ein Wechsel der Eingangsströme

und nach Einstellen des Gleichgewichtes resultiert ein neuer Zustand der Schaltung wie nach Abbildung 3.7, jedoch mit vertauschten Indizes von I_D und V_{DS} . Das Transientenverhalten dieses Überganges unterteilt sich in drei Abschnitte; im ersten Abschnitt erfolgt eine Änderung von $V_{GS(MS1)}$, da $V_{DS(MS1)}$ vorerst durch $V_{DS(MF1)}$ bestimmt bleibt. Die Änderung bewirkt einen neuen Arbeitspunkt von M_{S1} und die Kennlinie in Abbildung 3.7 verschiebt sich entsprechend. Dieser Vorgang ist in Abbildung 3.8 verdeutlicht. Die Spannung $V_{DS(MS2)}$ muss zur Einhaltung

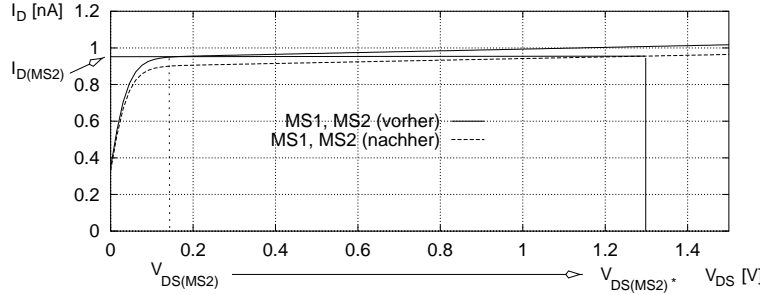


Abbildung 3.8: Verschiebung der Kennlinie von M_{S1} und M_{S2}

des Gleichgewichts den neu definierten Punkt der Kennlinie erreichen und steigt daher an. Die Anstiegs-Geschwindigkeit der Drain-Source-Spannung von M_{S2} bestimmt die Dauer des ersten Abschnitts. $V_{DS(MS2)}$ erreicht nach einer Weile eine bestimmte Höhe, ab der ein maßgeblicher Drain-Strom in M_{F2} fließen kann. Die Gate-Source-Spannung des M_{F2} steigt schließlich auf denselben Betrag wie $V_{GS(MF1)}$ an. An dieser Stelle findet der Wechsel der Ausgangssignale des WTA-Moduls statt, da beide M_{F1} und M_{F2} kurzzeitig den gleichen Drain-Strom führen. Im dritten Abschnitt verhindert der beginnende Stromfluss in M_{F2} den Stromfluss in M_{F1} , da die Beziehung $I_{D(MF1)} + I_{D(MF2)} = I_{D(MSRC)}$ gilt. Die Drain-Source-Spannung von M_{S1} muss einen neuen Arbeitspunkt einnehmen. Die Absink-Geschwindigkeit der Drain-Source-Spannung von M_{S1} bestimmt die Dauer des dritten Abschnitts. Es folgt eine Abschätzung der zeitlichen Dauer der Abschnitte.

Das Drain-Potential des M_{S2} verändert sich aufgrund des neu eingestellten Drain-Stromes $I_{D(MS2)}$ als Resultat des Absinkens von $V_{GS(MS2)}$ und steigt an. Dieser Vorgang lässt sich durch die Umformung der Schaltung von Abbildung 3.9 a) in ein Kleinsignal-Ersatzschaltbild von Teilbild b) erläutern. Nach Vereinfachung des Ersatzschaltbildes entsteht der RC-Kreis von Teilbild c), wobei

$$i = I_{ph2} - (v_{gs(MS2)} g_{m(MS2)}), \quad (3.7)$$

$$R = (g_{ds(MS2)} + g_j)^{-1} \quad \text{und} \quad (3.8)$$

$$C = C_j + C_{Dj(MS2)} + C_{G(MF2)} \quad (3.9)$$

ist. Der Strom i ist die Differenz des Photostromes und des neu definierten Drain-Stromes von M_{S2} , da durch das jetzt eingestellte Gate-Potential von M_{S1} ein geringerer Drain-Strom in M_{S2} fließt. Der resultierende Strom ermöglicht das Umladen der Kapazitäten des Ersatzschaltbildes und die Spannung $v(t)$ folgt der Beziehung

$$v(t) = \frac{1}{C} \int (i_{ph2} - v_{gs(MS2)} g_{m(MS2)}) dt. \quad (3.10)$$

Die Geschwindigkeit des Umladens ist zu der Differenz zwischen Drain-Strom und Photostrom proportional und zu der umzuladenden Kapazität umgekehrt proportional. Die Differenz zwischen Photostrom und nominalen Drain-Strom wird durch die Eingangs-Signale des WTA-

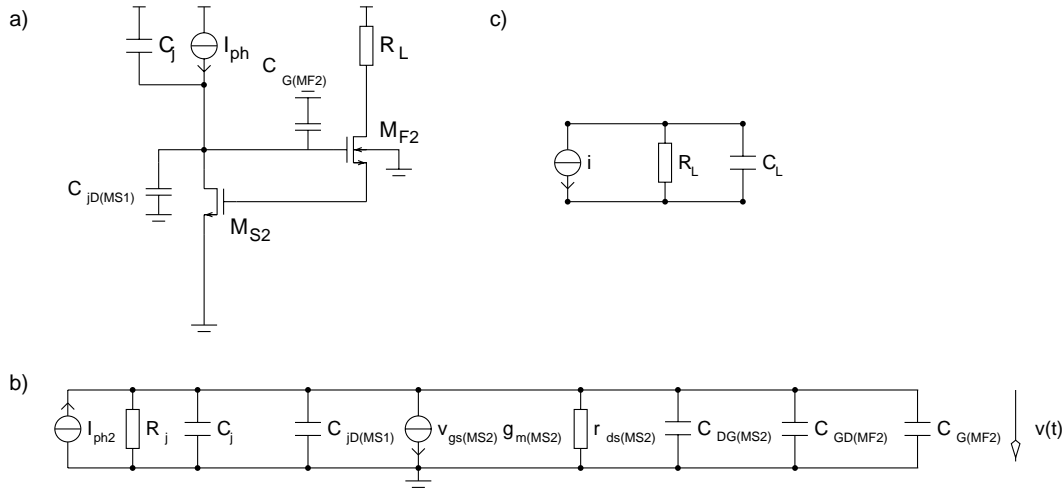


Abbildung 3.9: Ersatzschaltbild zur Berechnung der RC-Konstante

Moduls vorgegeben und kann bei entsprechenden Fokalfunkt-Formen und den daraus resultierenden geringen Photostrom-Unterschiede weniger als 1 pA betragen. Die Sperrschichtkapazität C_j des Photodetektors ist nach Gleichung 2.9

$$C_j = A \sqrt{\frac{\epsilon_{si} e N_A N_D}{2(N_A + N_D)(V_D - V)}}. \quad (3.11)$$

Bei Annahme einer Fläche einer einzelnen Photodiode eines Photodetektorstreifens von $A = 7.35 \mu\text{m} \times 15.85 \mu\text{m}$ ergibt sich unter Verwendung von [AMS 0.6 I] bei einer Sperrspannung von 2 V eine Sperrschichtkapazität von ca. 120 fF. Summation der 19 Photodioden eines Photodetektorstreifens ergibt $C_j \approx 2.28 \text{ pF}$, und unter Annahme einer Strom-Differenz von $I_{ph2} - I_{D(MS2)} = I_{ph2} - I_{ph1} = 1 \text{ pA}$ und Vernachlässigung der Kapazitäten $C_{Dj(MS2)}$ und $C_{G(MF2)}$ resultiert zum Einstellen einer um 1 V erhöhten Drain-Spannung eine Zeitdauer von 2.28 s. Diese Zeitdauer zum Einstellen der Drain-Source-Spannung von M_{S2} ist nicht zu vereinbaren mit der Zeit-Spezifikation des Projektes, die eine Wiederholrate von 1 kHz fordert. Da in der Anwendung Photoströme in der Höhe einiger pA und Differenzen zweier benachbarter Ströme mit einem Unterschied von nicht mehr als 20%, bei schlechter Fokalfunkt-Qualität sogar weniger als 10% auftreten können, ist diese Zeitdauer erheblich zu verkürzen. Die Hauptursache der langen Zeitdauer von 2.28 s ist der große Wert der Photodetektor-Sperrschichtkapazität. Der Wert selbst ist jedoch nicht zu verringern, da die Fläche eines Photodetektors proportional dem generierbaren Photostrom ist und eine Verkleinerung des Detektors eine Reduzierung des Photostromes implizieren würde. Eine Beschleunigung des transienten Verhaltens kann nur über die Reduzierung der von den Drain-Knoten der M_{Sn} "sichtbaren" Sperrschicht-Kapazitäten der Photodetektoren erfolgen.

Beschleunigung durch Entkopplung der Photodetektoren-Sperrschichtkapazität

Eine Abhilfe der zeitlichen Problematik der WTA-Schaltung ist die Verwendung einer Kaskode-Konfiguration, die die Spannung über der Sperrschicht-Kapazität der Photodetektoren von dem Potential der Drain-Knoten der M_{Sn} entkoppelt. Abbildung 3.10 stellt die modifizierte Schaltung dar. Eine Kaskode hat im Allgemeinen die Funktion, die Auswirkungen großer Ausgangsspannungs-Änderungen auf die Eingangs-Spannung einer Schaltung zu reduzieren. Besonders für das Frequenzverhalten einer Verstärker-Schaltung ist es wichtig, eine Entkopplung von Ein-

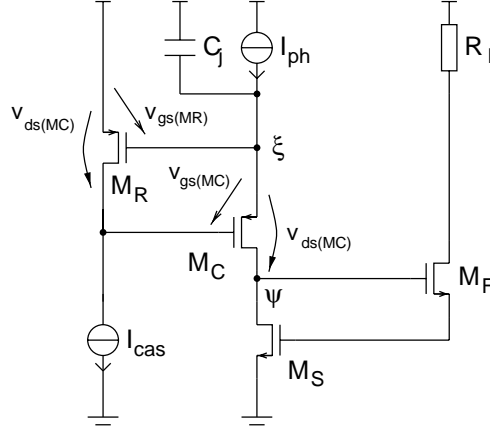


Abbildung 3.10: Berechnung der regulierten Kaskode. Eine große Änderung des Potentials des Knoten Ψ erzeugt eine kleine Änderung des Potentials des Knotens ξ .

und Ausgang zu realisieren, da sonst eine scheinbare Vergrößerung der Eingangskapazität¹¹ verursacht wird. Das Frequenzverhalten der Schaltung verbessert sich, indem durch die Kaskode die Änderung des Potentials des Knotens ξ als Folge einer Änderung des Potentials des Knotens Ψ reduziert wird. Die einfache Kaskode benutzt einen einzelnen Transistor M_C , dessen Gate ein konstantes Potential erhält. Der Effekt kann durch den Aufbau einer regulierten Kaskode unter Hinzufügen eines weiteren Transistors verstärkt werden. Das Potential des Knotens ξ ist definiert durch die Gate-Source-Spannung $v_{gs(MR)}$ des M_R . Die beiden Gate-Source-Spannungen $v_{gs(MR)}$ und $v_{gs(MC)}$ berechnen sich unter Vernachlässigung des Body-Effekts als

$$v_{gs(MC)} = \ln \frac{I_{ph}}{I_0 S (1 + \lambda(v(\Psi) - v_{gs(MR)}))} n V_t \quad \text{und} \quad (3.12)$$

$$v_{gs(MR)} = \ln \frac{I_{cas}}{I_0 S (1 + \lambda(v_{gs(MR)} + v_{gs(MC)}))} n V_t. \quad (3.13)$$

Ein totales Differential bestimmt die Änderung von $v(\xi)$ nach $v(\Psi)$:

$$\begin{aligned} \frac{dv(\xi)}{dv(\Psi)} &= \frac{dv_{gs(MR)}}{dv_{gs(MC)}} \cdot \frac{dv_{gs(MC)}}{dv(\Psi)} \\ &= \frac{\lambda^2 n^2 V_t^2}{(1 + \lambda(v_{gs(MR)} + v_{gs(MC)}) n V_t) (1 + \lambda(v(\Psi) - v_{gs(MR)}))} \\ &\approx \lambda^2 n^2 V_t^2 \end{aligned} \quad (3.14)$$

Die Sperrschicht-Kapazität C_j erscheint für den Knoten Ψ als $\lambda^2 n^2 V_t^2 \cdot C_j$. Nach Gleichung 3.9 ist die parasitäre Kapazität des Knotens Ψ daher

$$\begin{aligned} C &= \lambda^2 n^2 V_t^2 \cdot C_j + C_{Dj(MS)} + C_{G(MF)} + C_{Dj(MC)} \\ &\approx C_{Dj(MS)} + C_{G(MF)} + C_{Dj(MC)}, \end{aligned} \quad (3.15)$$

da $\lambda \ll 1$. C bestimmt sich zum Hauptteil aus den Sperrschicht-Kapazitäten der Drain-Gebiete von M_{S2} und M_{C2} und der Gate-Kapazität von M_{F2} , die sich nach [AMS 0.6 I] zu $C_{jD(MS2)} = 3.17 \text{ fF}$, $C_{jD(MC2)} = 2.85 \text{ fF}$ und $C_{G(MF2)} = 43.0 \text{ fF}$ berechnen, und beträgt ca. 49 fF. Die

¹¹Dieser Effekt wird auch als Miller-Effekt bezeichnet

Zeitverzögerung zur Änderung des Drain-Potentials von M_{S2} um 1 V beträgt bei $C \approx 49 \text{ fF}$ und $\Delta I_{ph} \approx 1 \text{ pA}$ ca. $\tau \approx 49 \text{ ms}$. Die errechnete Zeitdauer wird benötigt, um das Drain-Potential des M_S der “gewinnenden” Zelle bis zu einem Wert anwachsen zu lassen, ab dem ein Stromfluss in M_F etabliert wird und als Ausgangsstrom I_{out} verwendet werden kann.

Beschleunigung durch Initialisierung der Potentiale

Der transiente Wechsel der Ausgangsströme nach Wechsel der Eingangssignale findet wie oben beschrieben in drei Abschnitten statt; im ersten Abschnitt erfolgt die Änderung der Drain-Source-Spannung von M_{S2} , im zweiten der Wechsel der Ausgangsströme und im dritten die Änderung der Drain-Source-Spannung von M_{S1} . Dieser Vorgang kann mit Hilfe einer Initialisierung der Anfangsbedingungen beschleunigt werden, indem der erste Abschnitt durch ein erzwungenes Angleichen der Drain-Source-Spannungen der M_{Sn} entfällt. Dazu werden nach Aktivierung des Signals V_{res} zum Zeitpunkt t_0 die Drain-Knoten der M_{Sn} durch die Reset-MOS-FET M_{In} mit einem willkürlichen Potential V_{init} verbunden. Nach einer bestimmten Zeitdauer, in der die Schaltung den Zustand des Gleichgewichts einnehmen kann, egalisieren sich die Drain-Potentiale, und die Photoströme, die während der Initialisierungsphase generiert werden, fließen über die Reset-MOS-FET ab. In Abbildung 3.11 ist die erweiterte Architektur der WTA-Zellen verdeutlicht. Die Potentiale der Knoten Ψ_1 und Ψ_2 verändern sich zu V_{init} , da die Photoströme über

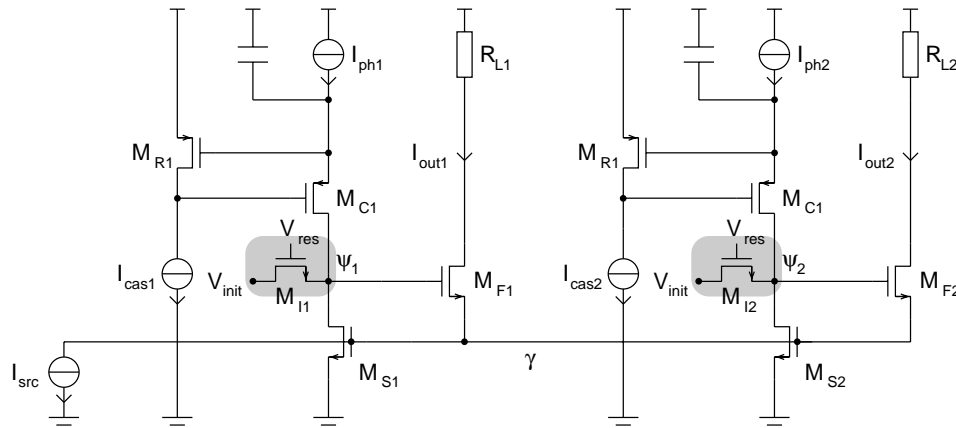


Abbildung 3.11: Initialisierung der Drain-Potentiale der M_{Sn}

die Reset-Transistoren, die im Vergleich zu allen anderen am Knoten befindlichen Widerständen niederohmiger sind, abfließen können. Der Strom I_{src} verteilt sich auf die Drain-Ströme der M_{Fn} , so dass die Gate-Source-Spannungen $V_{GS(MFn)}$ definiert werden, die als Differenz zu V_{init} das Potential des Knotens γ bestimmen. Die Drain-Ströme der M_{Sn} werden durch die Gate-Source- und Drain-Source Spannungen $V_{DS(MSn)}$ und $V_{GS(MSn)}$ definiert und durch die Spannungsquelle von V_{init} geliefert. Die Beträge der Ströme sind sehr klein, da die verbleibende Gate-Source-Spannung der M_{Sn} durch den exponentiellen Zusammenhang nach Gleichung 2.17 einen Abfall der Drain-Ströme um mehrere Dekaden verursachen.

Ausgehend von diesem Zustand der Schaltung wird über das Deaktivieren der Reset-Transistoren zu einem Zeitpunkt t_1 eine transiente Veränderung der Potentiale initiiert. Abbildung 3.12 verdeutlicht die ab diesem Zeitpunkt bestehende Konstellation der Schaltung. Ab dem Zeitpunkt t_1 sei $I_{ph1} > I_{ph2}$. Die Drain-Ströme der M_{Sn} sind aufgrund des niedrigen Potentials des Knotens γ kurz nach t_1 zu vernachlässigen. Die Photoströme I_{ph1} und I_{ph2} können daher nur in die parasitären Kapazitäten C_{Ψ_1} und C_{Ψ_2} abfließen und erhöhen die über den Kapazitäten abfallenden Spannungen und somit die Potentiale der Knoten Ψ_1 und Ψ_2 . Da die Gate-Source-Spannungen

und nach Definition der Bedingung $I_{out2}/I_{out1} = 0.1$ zur Bestimmung eines Zeitpunktes t_2 folgt

$$I_0 \operatorname{Se}^{\frac{V(\Psi_2)(t)-V(\gamma)(t)}{n V_t}} = 0.1 I_0 \operatorname{Se}^{\frac{V(\Psi_1)(t)-V(\gamma)(t)}{n V_t}}. \quad (3.20)$$

Nach Einsetzen der Gleichungen 3.17 für die Werte von $V(\Psi_1)$ und $V(\Psi_2)$ und Gleichsetzen der Anfangspotentiale der Knoten Ψ_1 und Ψ_2 mit $V(\Psi_1)(t_1) = V(\Psi_2)(t_1)$ ergibt sich

$$t_2 - t_1 = -\ln(0.1) n V_t \frac{C_\Psi}{\Delta I_{ph} - \int_{t_1}^{t_2} I_{D(MS2)}(t) - I_{D(MS1)}(t) dt}. \quad (3.21)$$

Die Drain-Ströme der M_{Sn} sind in erster Näherung gleich, da die Gate-Source Potentiale identisch und die Drain-Source-Potentiale bis zum Zeitpunkt t_2 annähernd gleich sind. Aus Gleichung 3.21 folgt für $I_{D(MS2)} = I_{D(MS1)}$ und unter Annahme von $C_\Psi \approx 49 fF$ und $\Delta I_{ph} \approx 1 pA$ eine Zeitdauer von $t_2 - t_1 = 4.2 ms$, um ein Verhältnis der Ausgangsströme von $I_{out2}/I_{out1} = 0.1$ zu erreichen.

Abbildung 3.13 stellt eine Simulation der Schaltung von Abbildung 3.11 unter Betrachtung der Spannungen $V(\Psi_1)$, $V(\Psi_2)$ und $V(\gamma)$ sowie der beiden Ströme I_{out1} und I_{out2} dar. Die Reset-MOS-FET sind bis zum Zeitpunkt t_1 aktiviert. Nach der Deaktivierung durch $V_{set} = 0$ bei t_1 erfolgt beginnend bei $V_{init} = 0.7 V$ das Aufladen der C_{Ψ_n} . Das Potential des Knotens Ψ_1 steigt schneller als das des Knotens Ψ_2 , da mehr Photostrom zum Umladen der Kapazität des Knotens vorhanden ist. Das Potential des Knotens γ steigt ebenfalls an, da es durch die Gate-Source-Spannungen der M_{Fn} mit den Potentialen der Knoten Ψ_1 und Ψ_2 verbunden ist. Die Differenz der Spannung $V(\Psi_1) - V(\gamma)$ verursacht den Strom I_{out1} , die Differenz $V(\Psi_2) - V(\gamma)$ den Strom I_{out2} . Die Differenz $V(\Psi_1) - V(\Psi_2)$ steigt stetig an, obwohl $V(\Psi_1)$ und $V(\Psi_2)$ ein Maximum erreichen. In diesem Moment fließt ein nennenswerter Drain-Strom in den M_{Sn} , der das weitere Aufladen der C_{Ψ_n} verhindert. Die Kapazität C_{Ψ_2} wird im weiteren Verlauf durch die Differenz der Photoströme wieder entladen, so dass das Potential des Knotens Ψ_2 sinkt.

Die Ströme I_{out1} und I_{out2} betragen beide $0.5 I_{src}$ vor dem Zeitpunkt t_1 . Ab t_1 steigt I_{out1} an und I_{out2} fällt ab. Die Summe der beiden Ströme beträgt I_{src} . Nach einiger Zeit stellt sich ein Strom $I_{out1} = I_{src}$ und $I_{out2} \rightarrow 0$ ein, da sich die Arbeitspunkte der M_{Sn} entsprechend Abbildung 3.5 eingestellt haben. Die Simulation benutzt als Modelle der Transistoren M_S , M_F , M_R , M_C und M_I die Gleichung 2.19, die durch die analoge Hardware-Beschreibungssprache *SpectreHDL* beschrieben wird. Die Werte von n , S , I_0 und λ betragen jeweils 1.5, 0.5, 1 fA und $0.05 V^{-1}$. Die Werte der Eingangsströme betragen $I_{ph1} = 11 pA$, $I_{ph2} = 10 pA$, I_{src} beträgt 60 nA.

Beschleunigung durch positive Rückkopplung

Die benötigte Zeit der WTA-Zellen zum Einstellen der Ausgangssignale entspricht bei einer Differenz der Photoströme von 1 pA nach Gleichung 3.21 ca. 4.2 ms. Im praktischen Betrieb des HSSX kann es vorkommen, dass die Photoströme bei einer sehr geringen Intensität der Fokalfunkte Werte von nicht mehr als einem pA annehmen und eine Differenz der Photoströme benachbarter Photodetektor-Streifen von ca. 100 fA entsteht. Unter Ansetzen der Gleichung 3.21 beträgt die Zeitdauer zum Einstellen der Ausgangsströme bei einer Photostromdifferenz von 100 fA ca. 42 ms. Diese Zeitdauer lässt sich durch Einführung einer positiven Rückkopplung der Ausgangsströme in die Eingangsströme der WTA-Zellen verringern. Die sogenannten *Feedback*-Ströme unterstützen die Photoströme im Aufladen der Knotenkapazitäten C_{Ψ_n} , so dass sich die Potentiale der Knoten Ψ_1 und Ψ_2 schneller erhöhen. Dabei werden die Beträge der Feedback-Ströme auf Stromhöhen reduziert, die ungefähr der Höhe der Eingangssignale entsprechen. Im Folgenden wird das Konzept der Rückkopplung erläutert, sowie Maßnahmen zur Regulierung der Verstärkung und zum Abschalten der Rückkopplung vorgestellt.

der Kanallängen-Modulation beträgt das Verhältnis der Spiegelströme

$$\frac{I_{out}}{I_{fb}} = \frac{I_0 S e^{\frac{V_{GS(MM)}}{n V_t}}}{I_0 S e^{\frac{V_{GS(MFB)} + V_{SB(MFB)}}{n V_t}} e^{-\frac{V_{BS(MFB)}}{V_t}}} = e^{\frac{V_{SB(MFB)}(n-1)}{n V_T}}. \quad (3.23)$$

Das Einstellen einer Bulk-Source-Spannung von z.B. 0.6 V mit $n = 1.5$ und $V_t = 0.025$ V verursacht ein Verhältnis $I_{fb} = \frac{1}{A} I_{out}$ mit $A \approx 3 \cdot 10^3$.

Die Übertragungsfunktion des Systems unter Rückkopplung und Anliegen einer Photostromdifferenz lässt durch die zeitliche Abhängigkeit der Ausgangsströme von der Differenz der Potentiale der Knoten Ψ_{nn} beschreiben. Die Ausgangsströme folgen den Bedingungen $I_{out1} + I_{out2} = I_{src}$ und

$$\frac{I_{out1}}{I_{out2}} = e^{\frac{\Delta V_{GS(MF)}}{n V_t}}. \quad (3.24)$$

Die Ströme I_{out1} und I_{out2} werden explizit als Funktion der Potentialdifferenz der Knoten Ψ_n formuliert und lauten

$$\begin{aligned} I_{out1} &= I_{src} \frac{1}{1 + e^{-\Delta V_{GS(MF)}}} \quad \text{und} \\ I_{out2} &= I_{src} \frac{1}{1 + e^{\Delta V_{GS(MF)}}}. \end{aligned} \quad (3.25)$$

Die Rückkopplung erzeugt Feedback-Ströme, deren Differenz zu einem Aufladen der Kapazitäten C_Ψ beitragen. Unter Vernachlässigung einer Differenz der Drain-Ströme der M_{Sn} resultiert

$$\Delta V(\Psi)(t) = \frac{1}{C_\Psi} \int_{t_1}^{t_2} \Delta I_{ph} + \frac{1}{A} \cdot (I_{out1} - I_{out2}) dt, \quad (3.26)$$

und unter Einsetzen von 3.25 lautet die Übertragungsfunktion

$$\Delta V(\Psi)(t) = \frac{1}{C_\Psi} \int_{t_1}^{t_2} \Delta I_{ph} + \frac{1}{A} I_{src} \frac{1 - e^{-\Delta V(\Psi)(t)/n V_t}}{1 + e^{-\Delta V(\Psi)(t)/n V_t}} dt. \quad (3.27)$$

Diese Differentialgleichung ist nicht explizit lösbar, kann jedoch unter numerischer Differentiation berechnet werden. Die Funktionalität der Rückkopplung wird daher durch Simulation der Schaltung aus Abbildung 3.14 verifiziert. Der Simulation liegen die Modelle und Parameter des vorangehenden Abschnitts zugrunde. Abbildung 3.15 zeigt das Einstellen der Ausgangsströme unter Anliegen einer Photostrom-Differenz von 100 fA und einem Untersetzungsverhältnis von $A = 1 \cdot 10^4$ und $A \rightarrow \infty$. $t_2 - t_1$ entspricht der im vorangehenden Abschnitt berechneten Zeitdauer von 42 ms, um die Ausgangsströme bei Anliegen einer Photostrom-Differenz von $\Delta I_{ph} = 100$ fA mit $A \rightarrow \infty$ in einem Verhältnis von $I_{out2}/I_{out1} = 0.1$ zu etablieren. Die Zeitdauer $t_2^* - t_1$ von 0.95 ms verdeutlicht die Beschleunigung des Vorganges unter $A = 3 \cdot 10^3$ und stimmt mit numerischen Berechnungen von Gleichung 3.27 überein.

Die Rückkopplung wird verwendet, um bei Photostrom-Differenzen von unter 1 pA in einer mit den Spezifikationen zu vereinbarenden Zeitdauer die Ausgangssignale zu etablieren. Die Untersetzung der Stromspiegel wird derart eingestellt, dass die Rückkopplung für $\Delta I_{ph} < 1$ pA dominant und für $\Delta I_{ph} > 1$ pA rezitiv ist. Bei einer Stromstärke der Quelle $I_{src} = 60$ nA und einer Anzahl von 20 WTA-Zellen beträgt die einzustellende Spannung des Bulk-Knotens der M_{FBn} ca. 520 mV. Dadurch wird ein Untersetzungsverhältnis von $A = 1 \cdot 10^3$ erzeugt und die Feedback-Ströme betragen nominal ca. 6 pA.

In einer definierten Integrationsdauer, in der die Ladung Q in einem Kondensator gespeichert wird, beträgt die Unsicherheit der Spannung des Kondensators C nach [Decker 98] daher

$$\overline{v_{n1}^2} = (e n_n \frac{1}{C})^2 = e \frac{Q}{C^2}. \quad (3.29)$$

In dem *HSSX* findet die Integration der Photoströme in den parasitären Kapazitäten der Knoten Ψ statt. Die auf Rauschen und Offset-Schwankungen sensitiv reagierenden Elemente einer WTA-Zelle sind die Flow-Transistoren M_{Fn} . Die Differenz der Gate-Source-Spannungen der M_{Fn} bestimmt das Ausgangssignal der Zelle und somit erfolgt die Berechnung der äquivalenten Rauschspannung durch die Betrachtung der Potentiale der Knoten Ψ .

Zum Ende der Initialisierungs-Phase beträgt die Spannung der Knoten Ψ V_{init} . Nach Abschalten der Reset-MOS-FET laden der Photostrom und der Sperrstrom der Photodioden die parasitären Kapazitäten der Knoten Ψ auf. Das statistische Überqueren der Ladungsträger des pn-Überganges der Photodetektoren verursacht für den Photostrom und den Sperrstrom einen Anteil von Rauschstrom, der innerhalb der Integrations-Phase eine Verteilung der tatsächlichen Potentiale der Knoten Ψ bewirkt. Unter Ansetzen von Gleichung 3.29 und der Ladungsmenge Q , die durch Photostrom und Sperrstrom der Photodetektoren geliefert wird als

$$Q = \int_0^t I_{ph} + I_s d\tau, \quad (3.30)$$

ergibt sich zum Beispiel unter Verwendung von $I_{ph} = 10 pA$, $I_s = 50 fA$ und $C_\Psi = 49 fF$ innerhalb einer Zeitdauer von $t = 1 ms$ eine äquivalenten Rausch-Spannung der Knoten Ψ von $v_{n1} \approx 800 \mu V_{rms}$.

Zu dieser Unsicherheit der Spannung über den Integrations-Kondensatoren C_Ψ muss eine während der Resetphase entstehende Rauschspannung an den Knoten Ψ addiert werden. Diese Rauschspannung, als *kTC noise*¹² bezeichnet, wird nach Beendigung der Resetphase "eingefroren" und berechnet sich durch Integration der Rauschspannung der Knoten Ψ in Abhängigkeit des Rauschstromes und der Impedanz der Knoten im Frequenzbereich.

Die spektrale Impedanz beträgt unter Vernachlässigung des Kanalwiderstandes der Transistoren M_C (siehe Abbildung 3.11) aufgrund von Degeneration durch den hohen Widerstand der Stromquellen I_{ph}

$$Z(\omega) = (j\omega C_\Psi + g_{ds(MI)} + g_{ds(MS)})^{-1}. \quad (3.31)$$

Die kTC-Rauschspannung der Knoten Ψ berechnet sich durch Integration des Produktes aus spektralem Anteil des Rauschstromes und der frequenzabhängigen Impedanz:

$$\overline{v_{n2}^2} = \int_0^\infty \frac{2 e (I_{ph} + I_s)}{|j2\pi f C_\Psi + g_{ds(MI)} + g_{ds(MS)}|^2} df. \quad (3.32)$$

Die dritte Unsicherheit der Spannung über den Integrations-Kondensatoren C_Ψ resultiert aus einer Verteilung der Ladungsmenge, die durch den Spannungssprung des Gate-Potentials der M_I in den Kanal der MOS-FET injiziert wird. Die Menge der Elektronen n_i , die durch dieses sog. *Clock-Feedthrough* der Transistoren in den Kanal eingebracht wird, berechnet sich unter Betrachtung der Spannungsänderung $\Delta V(\Psi)$ nach [Allen 87d] von

$$\Delta V = -\frac{C_{GS}}{C_{GS} + C_\Psi} (V_{init} + V_{THn}), \quad (3.33)$$

¹²Thermisch induziertes Rauschen einer Spannung über einem Kondensator

mit der Kapazität C_{GS} als Überlappkapazität der M_I zwischen Gate und Source, V_{init} als Initialisierungsspannung und V_{THn} als Threshold-Spannung der M_I , zu

$$n_i = \frac{1}{e} C_\Psi \Delta V. \quad (3.34)$$

Unter Ansetzen der Poisson-Verteilung entspricht die Anzahl der Rauschelektronen, die in den Kanal gelangen, der Wurzel aus n_i . Damit lässt sich die Unsicherheit der Potentiale der Knoten Ψ durch das *Clock-Feedthrough* der M_I quantifizieren als

$$\overline{v_{n3}^2} = \frac{1}{C_\Psi^2} (\sqrt{n_i} e)^2 = \frac{1}{C_\Psi} e \Delta V. \quad (3.35)$$

Die vierte und fünfte Unsicherheit der Spannung kann durch die in die Eingangsspannung der Transistoren M_F transformierte Unsicherheit der Kanalströme der M_F berechnet werden. Sie beinhaltet die Auswirkung des Funkel-Rauschens und des thermischen Rauschens (vergl. Abschnitt 2.3) mit den Anteilen

$$\begin{aligned} \overline{v_{n4}^2} &= \int_0^{f_2} \frac{KF}{2 f C_{ox} WL K'} df \quad \text{und} \\ \overline{v_{n5}^2} &= \int_0^{f_2} \frac{8 k T}{3 g_m} df. \end{aligned} \quad (3.36)$$

Die Bandbreite der Übertragungsfunktion der M_{Fn} berechnet sich durch eine Kleinsignalanalyse der Schaltelemente I_{src} , M_{Fn} , M_{Mn} und der parasitären Kapazität C_κ der Knoten κ zu

$$f_2 = 1.57 \cdot f_1 = 1.57 \cdot \frac{1}{2\pi} \frac{g_{m(MM)} + g_{ds(MF)}}{C_\kappa}. \quad (3.37)$$

Die Relation $f_2 = 1.57 f_1$ in Gleichung 3.37 berücksichtigt nach [Grey 84b] das Frequenzverhalten einer Übertragungsfunktion mit einem Pol an der Stelle f_1 .

Aufgrund der Unabhängigkeiten der verschiedenen Rauschspannungsquellen $\overline{v_{ni}^2}$ voneinander ergibt die Summation der einzelnen Quellen den Rauschbetrag der äquivalenten Rauschspannungsquelle $\overline{v_n^2}$ von Abbildung 3.16 mit

$$\overline{v_n^2} = \sum_{i=1}^5 \overline{v_{ni}^2} = \overline{v_{np}^2}(I_{ph}) + \overline{v_{nc}^2}, \quad (3.38)$$

wobei die Variablen $\overline{v_{np}^2}(I_{ph})$ den von der Höhe der durch den Photostrom erzeugten Ladung der Kondensatoren C_Ψ abhängigen Rauschanteil und $\overline{v_{nc}^2}$ den unabhängigen Rauschanteil repräsentieren.

Die Spannung, die sich durch das Aufladen der Kondensatoren durch Photostrom und Sperrstrom in der definierten Zeitdauer ergibt, beträgt unter Annahme der oben verwendeten Werte für I_{ph} , I_s , t und C_Ψ

$$V = \frac{1}{C_\Psi} \int_0^t I_{ph} + I_s d\tau \approx 201 mV. \quad (3.39)$$

Bei Annahme eines Photostromes von $I_{ph} = 10 pA$ und unter numerischer Berechnung der Rauschspannungen $\overline{v_{ni}^2}$ mit Verwendung typischer Parameter bestimmt sich das Signal-Rauschverhältnis der Spannung der Knoten Ψ zum Beispiel zu

$$SN = \frac{V}{\sqrt{\overline{v_n^2}}} = \frac{201 mV}{2.7 mV_{rms}} \approx 75. \quad (3.40)$$

Der Anteil der Rauschspannung von der nominalen Spannung beträgt bei einem Photostrom von 10 pA demnach ca. 1.3%. Bei annähernd gleicher Stromhöhe der Eingangsströme von zwei benachbarten Zellen ist es möglich, dass die Differenz der Potentiale der beiden Knoten Ψ_1 und Ψ_2 einen geringeren Wert annimmt als die oben berechneten $v_n \approx 2.7 mV_{rms}$. In diesem Fall ist die Entscheidung, welche der beiden Zellen "gewinnen" wird, statistisch, und es entsteht Rauschen.

Durch die Unterteilung der Rauschspannung in einen von der Höhe des Photostromes abhängigen und einen unabhängigen Anteil folgt das Signal-Rausch-Verhältnis zuerst einem proportionalen Zusammenhang und geht anschließend aufgrund der Dominanz des Anteils von v_{n1}^2 in die Abhängigkeit von einer Quadratwurzel über. Die Vervierfachung der Ladungsmenge Q durch Erhöhung des Photostromes oder Verlängerung der Integrationsdauer ermöglicht eine Verdoppelung des Signal-Rauschabstandes.

Im Feedback-Betrieb, in dem die Ausgangsströme der M_F über die Spiegelkonfiguration aus M_M und M_{FB} in den Eingang der WTA-Zelle zurückgeführt wird, verringert sich der Signal-Rausch-Abstand, da zusätzlich zu dem Photostrom und dem Sperrstrom der Photodiode ein Feedback-Strom mit Rauschanteil das Aufladen der Kondensatoren C_Ψ bewirkt. Nach den Ausführungen des vorangehenden Abschnitts folgt die Übertragungsfunktion im Feedbackbetrieb einer nichtlinearen Differentialgleichung, die explizit nicht lösbar ist, sondern nur numerisch berechnet werden kann. Die quantitative Beurteilung des Rauschens im Feedback-Betrieb erfolgt daher durch Messungen in Kapitel 6.

Mismatching resultiert während des Herstellungsprozesses eines integrierten Schaltkreises durch Unregelmäßigkeiten in der Erzeugung der einzelnen Strukturen. Dabei ist zwischen einer Fehl-übereinstimmung von Strukturen desselben Wafers und zwischen Parameter-Streuungen unter mehreren Wafern zu unterscheiden. Im Folgenden wird das Mismatching von Strukturen mit geringem Abstand auf der Wafer-Oberfläche behandelt.

Die Konsequenz von Mismatching lässt sich durch Annahme von unterschiedlichen Schwellspannungen zweier MOS-FET zur Generation eines identischen Drainstromes abschätzen. In [AMS 0.6 I] erfolgen Angaben über das Mismatching identischer MOS-FET, die an dieser Stelle zitiert werden. Das Mismatching der im 0.6- μ m-CMOS-Technologie realisierbaren NMOS-Transistoren wird definiert als

$$\sigma(\Delta V_{THn}) = \frac{15.5 mV}{\sqrt{WL}}. \quad (3.41)$$

NMOS-FET mit Strukturgrößen von z.B. $W/L = 6/4$ besitzen eine Standardabweichung der nominalen Schwellspannung von ca. 3.2 mV. Die Angabe ist für Abschätzungen des Mismatching zwischen MOS-FET, deren Arbeitspunkt in dem Subthresholdbereich liegt, zu übertragen, da der Ausdruck zur Berechnung des Drain-Stromes in diesem Fall unabhängig von der Schwellspannung ist. Es wird daher ein Offset der Gate-Source-Spannung angenommen, dessen Standardabweichung zwischen den identischen Strukturen durch den Betrag aus Gleichung 3.41 angenähert wird. Nach Gleichung 2.23 unterscheiden sich die Drain-Ströme zweier identischen NMOS-FET-Strukturen mit einem Gate-Source-Spannungs-Offset von 3.2 mV mit $n = 1.5$ und $V_t = 0.025$ um 8.9 %.

In Abbildung 3.16 sind Serien-Spannungsquellen V_{osn} eingezeichnet, die das Verhalten aller Mismatching-Quellen der WTA-Zellen repräsentieren sollen. Dabei gibt es drei Quellen von Mismatching:

- Unterschiedliches Clock-Feedthrough der M_{In} auf die Knoten Ψ_n
- Offset der Gate-Source-Spannung der M_{Fn}

- Offset der Gate-Source-Spannung der M_{FBn}

Es ist anzumerken, dass ein Offset der Gate-Source-Spannung der M_{Sn} keinen Einfluss auf die Funktionalität der Schaltung hat, da die Ströme und die Unterschiede der Ströme $I_{D(MSn)}$ zu Beginn der Integrationsphase zu vernachlässigen sind.

Durch Simulationen der Schaltung mit aktivierter und deaktivierter Rückkopplung konnte abgeschätzt werden, dass die Funktionalität des WTA-Moduls durch Mismatching beeinträchtigt wird. Unter Annahme einer Photostromdifferenz von 10% konnte in dem Betriebsmodus ohne Rückkopplung eine Offsetspannung von 30 mV angenommen werden, ab der die korrekte Funktion nicht mehr gegeben war. Die Schaltung ist in dieser Betriebsart robust gegen Mismatching. In dem Betriebsmodus mit Rückkopplung verhinderte eine Offset-Spannung von mehr als 5 mV die Funktionalität. Die Quantifizierung des tatsächlich vorhandenen Mismatchings erfolgt in Kapitel 6.

Der Ausgangs-Buffer des WTA-Moduls

Entsprechend den Ausführungen des Abschnitts 3.1 wird die Information des WTA-Moduls über die Photodiodenzeile- oder Spalte des höchsten Photostromes in ein Schieberegister geladen. Das Ausgangssignal des WTA-Moduls wird durch die Höhe des Stromes I_{out} gebildet. Es erfolgt eine Umwandlung des Stromes in ein digitales Spannungssignal, um die Information des Moduls in eine digitale Eingangsstruktur der Schieberegister laden zu können. Der Spannungspegel des Ausgangs-Buffers soll entsprechend der eingesetzten CMOS-Prozess-Technologie einen vollen Spannungsbereich umfassen, um die nachfolgende Eingangsstufe des Schieberegisters mit einem Signal ausreichenden Störabstandes versorgen zu können. Die Stärke des Buffers bestimmt die Zeitdauer, die zu einem sicheren Erreichen der Schaltpegel benötigt wird. Sie wird durch die W/L-Parameter der MOS-FETs des Ausgangs-Buffers bestimmt. Abbildung 3.17 a) zeigt die Struktur der Eingangsstufe einer Schieberegister-Zelle, um die Belastung des Ausgangs-Buffers einer WTA-Zelle zu erläutern. Über M_1 kann ein Spannungspegel V_{load} durch Aktivierung von V_{set} an den Inverter M_p / M_n übertragen werden. Die beiden MOS-FET M_1 und M_2 bilden einen Multiplexer. M_2 sei für die folgende Betrachtung im Cutoff-Bereich und wird daher nur durch die Source-Bulk-Kapazität $C_{jD(M2)}$ ersetzt. Der Body-Effekt aller MOS-FETs wird in der Betrachtung vernachlässigt. Der Widerstand $R_{DS(M1)}$ symbolisiert den Kanalwiderstand

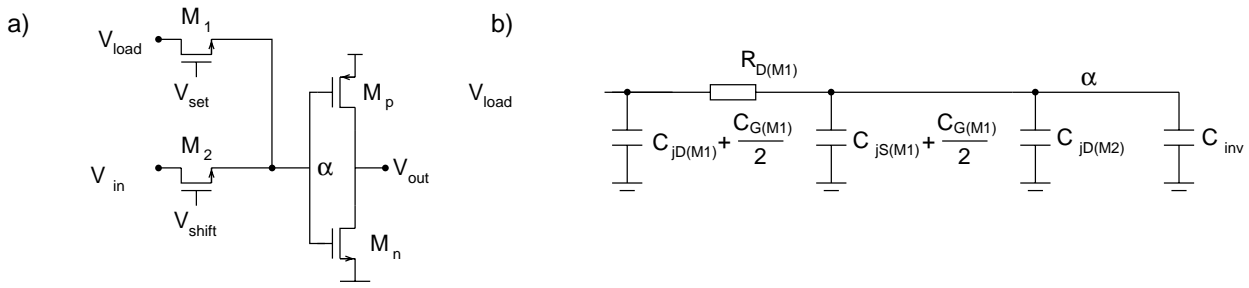


Abbildung 3.17: Über M_1 wird das Ausgangssignal des WTA-Moduls in die Schieberegister geladen

des Pass-Transistors M_1 und kann unter Annahme des linearen Betriebsbereiches des MOS-FET durch Gleichung 3.42 modelliert werden. Zur Näherung erster Ordnung werden der Term zweiter

Ordnung von $V_{DS(M1)}$ und der Faktor der Kanallängen-Modulation vernachlässigt.

$$R_{DS}^{-1} = \frac{\frac{\beta}{2} (V_{GS(M1)} - V_{TH(M1)}) V_{DS(M1)} - V_{DS(M1)}^2 (1 - \lambda V_{DS(M1)})}{V_{DS(M1)}} \approx \beta (V_{GS(M1)} - V_{TH(M1)}) \quad (3.42)$$

Der Kanalwiderstand ist durch den Faktor β , mit $\beta = S k$, durch Wahl von S einstellbar. In dem Projekt wird aufgrund der Vorgaben des Raumbedarfs eine minimale Transistorgröße gewählt, so dass ein Widerstand von einigen $k\Omega$ resultiert. Die Kapazität, die über $R_{DS(M1)}$ umgeladen wird, setzt sich aus den beiden Drain-Bulk- und Source-Bulk-Kapazitäten von M_1 , der Gate-Bulk-Kapazität von M_1 , der Drain-Bulk-Kapazität von M_2 , und der Eingangskapazität C_{inv} des Inverters zusammen, wobei die jeweils auftretenden Gate-Drain- oder Gate-Source-Kapazitäten vernachlässigt werden. Die Gesamt-Kapazität lässt sich unter Vorgabe der Prozessparameter der $0.6 \mu m$ -CMOS-Technologie abschätzen zu ca. 12.9 fF . Die Eingangsstufe des Schieberegisters kann folglich durch eine RC-Kette wie in Abbildung 3.17 b) modelliert werden.

Die Ausgangsstufe einer WTA-Zelle wird mit einem Klasse-A¹³-Verstärker realisiert, wobei der Eingangstransistor des Verstärkers als Ausgangstransistor eines Stromspiegels implementiert werden kann. Der Stromspiegel erhält als Eingangssignal das Stromsignal $I_{out,n}$ der WTA-Zelle. Diese Architektur ist in Abbildung 3.18 a) gezeigt. Der Strom I_{in} wird durch die Konfiguration

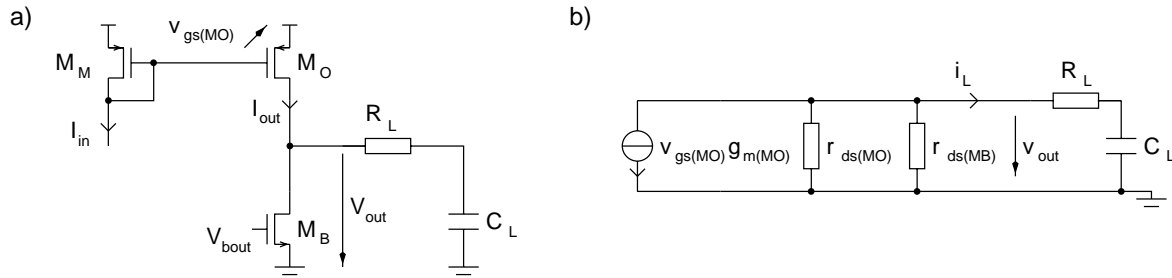


Abbildung 3.18: Die Last des WTA-Ausgangs-Buffers

M_M und M_O in den Strom I_O gespiegelt. M_B stellt aufgrund der Gate-Source-Spannung V_{bout} einen konstanten Strom zur Verfügung und zur Einhaltung der Beziehung $I_{D(MO)} = I_{D(MB)}$ stellt sich eine entsprechende Ausgangsspannung V_{out} ein, die für beide M_O und M_B die Lösung der Gleichung 2.19 ergibt. Dabei gilt $V_{DS(MO)} = V_{dd} - V_{out}$ und $V_{DS(MB)} = V_{out}$. Im Falle einer Änderung des Stromes I_{in} ergibt sich eine Änderung von V_{out} , die durch die Betrachtung von Teilbild 3.18 b) quantifiziert werden kann. Die Widerstände $r_{DS(MO)}$ und $r_{DS(MB)}$ liegen parallel zu $R_L + \frac{1}{j\omega C_L}$. $\frac{1}{j\omega C_L}$ ist in erster Näherung aufgrund des Spannungssprunges zu vernachlässigen. Der Strom $v_{gs(MO)} g_m(MO)$ verteilt sich auf die beiden Strompfade und ein größerer Teil durchfließt R_L , als Konsequenz der hohen Widerstandswerte von $r_{DS(MO)}$ und $r_{DS(MB)}$ gegenüber R_L , die entsprechen den Beträgen $r_{ds(MB)}^{-1} = \lambda_{MB} I_{D(MB)}$, $r_{ds(MO)}^{-1} = \lambda_{MO} I_{D(MB)}$ bei $\lambda = 0.05$ und $I_{D(MB)} \approx 100 \text{ nA}$ ca. $200 \text{ M}\Omega$ betragen. Daher gilt $i_L \approx v_{gs(MO)} g_m(MO)$ und das Auf- und Entladen der Kapazität C_L geschieht entsprechend der Formel

$$C_L \Delta v_{out} = i_L t. \quad (3.43)$$

Der Ausgangsstrom I_{out} für das Aufladen der Kapazität kann als Differenz des Ruhestromes $I_{D(MB)}$ und des Ausgangsstromes des Stromspiegels $I_{D(MO)}$ angesehen werden. Das Entladen

¹³ Ausgangsverstärker mit konstant fließendem Ruhestrom

findet in Näherung nur durch den Strom $I_{D(MB)}$ statt, da der Ausgangsstrom einer “verlierenden” WTA-Zelle sehr viel kleiner als $I_{D(MB)}$ ist. Bei Annahme eines Ausgangsstromes $I_{D(MO)}$ von ca. 100 nA und eines Ruhestromes $I_{D(MB)}$ von ca. 50 nA beträgt die Auf- und Entladezeit der Kapazität ca. 1.3 μ s, mit $C_L \approx 12.9$ fF und $\Delta v_{out} = 5$ V. Eine Verkürzung dieser Zeitdauer ist durch Erhöhung des Ruhestromes von M_B und des Signalstromes I_{in} möglich. Der Ruhestrom fließt jedoch in allen 10240 WTA-Zellen des *HSSX* und wird daher so gering wie möglich mit 50 nA gewählt, um diesen Teil des gesamten Stromaufkommen des *HSSX* in einem vertretbaren Rahmen zu halten.

Der Ausgangsspannungs-Bereich des Buffers lässt sich unter Gleichsetzen der Drainströme von M_O und M_B abschätzen. Die Eingangsgröße des Buffers ist die Gate-Source-Spannung von M_O und wird durch den Stromspiegel bei Anlegen eines Ausgangsstromes I_{out} der WTA-Zelle über die Gate-Source-Spannung des M_M definiert. Zur Berechnung der Ausgangsspannung seien die zwei Zustände einer WTA-Zelle betrachtet: die Zelle verursacht einen Ausgangsstrom von $I_{out} = I_{src}$ oder $I_{out} \rightarrow 0$. Die Gate-Source-Spannung von M_O ist in erster Näherung abhängig von dem Ausgangsstrom I_{out} , da $V_{GS(MO)} = V_{GS(MM)}(I_{out})$ gilt. Das Gleichsetzen der beiden Drain-Ströme von M_O und M_B

$$I_o S e^{\frac{V_{GS(MO)}}{n V_t}} (1 - e^{-\frac{V_{dd} - V_o}{V_t}}) (1 + \lambda_{MO} (V_{dd} - V_o)) = I_o S e^{\frac{V_{GS(MB)}}{n V_t}} (1 - e^{-\frac{V_o}{V_t}}) (1 + \lambda_{MB} (V_o)) \quad (3.44)$$

ergibt unter Annahme von $I_{src} = 2 I_{D(MB)}$ ¹⁴ die transzendente Gleichung

$$2(1 - e^{-\frac{V_{dd} - V_o}{V_t}}) (1 + \lambda_{MO} (V_{dd} - V_o)) = (1 - e^{-\frac{V_o}{V_t}}) (1 + \lambda_{MB} (V_o)), \quad (3.45)$$

die sich numerisch iterativ nach V_o lösen lässt. Unter Annahme der Werte $V_t = 0.025$ V, $V_{dd} = 5$ V und $\lambda = 0.3$ ergibt sich für $I_{out} = I_{src}$ eine Ausgangsspannung von $V_o = 4.44$ V, bei $I_{out} \rightarrow 0$ ein Wert von $V = 3.25$ nV $\rightarrow 0$ V. Der Wert $\lambda = 0.3$ entspricht dem Kanallängen-Modulationsfaktor eines MOS-FET mit der Eigenschaft $W/L \approx 2/1.2$. Eine höhere Ausgangsspannung ließe sich unter Reduzierung des Faktors λ durch Verlängerung der Kanallänge von M_O und M_B erreichen. Dies ist jedoch aufgrund der kritischen Flächenspezifikationen des *HSSX* nicht akzeptabel. Weiterhin ist die Maßnahme nicht erforderlich, da der Eingangsmultiplexer der Schieberegisterzelle einen NMOS-FET als Pass-Transistor verwendet. Da ein Pass-Transistor eine maximale Eingangsspannung von ca. 4.2 V durchschalten kann¹⁵, ist die Höhe der Ausgangsspannung von 4.44 V ausreichend.

3.2.4 Die dynamischen Schieberegister

Die während der Dauer einer Detektion erzeugten Ausgangspotentiale werden nach Beendigung der Detektion in Schieberegister-Zellen geladen. Nach den Ausführungen in dem vorangehenden Abschnitt erzeugt eine WTA-Zelle ein digitales Spannungssignal, dass durch einen Pass-Transistor in ein Schieberegister übertragen wird. Die Schieberegister werden nach Beendigung des Ladevorgangs über den zweiten Pass-Transistor des Eingangs-Multiplexers zu einer Kette zusammengeschaltet und die gespeicherten Informationen durch eine Daisy-Chain aller Schieberegister an den Kompressor übertragen. Abbildung 3.19 zeigt den schematischen Aufbau zweier WTA-Module und der zugehörigen Schieberegister. In der Abbildung sind die WTA-Module zur Bestimmung der X-Positionen der Fokalfpunkte einer Cluster-Zeile dargestellt. Die Schieberegister aller WTA-Module der Zeile sind miteinander verbunden, so dass sich die X-Positions-Daten von insgesamt 16 Clustern kaskadiert auslesen lassen. Im folgenden Abschnitt wird der Aufbau einer Schieberegister-Zelle erläutert.

¹⁴Diese Annahme definiert $e^{\frac{V_{GS(MO)}}{n V_t}} = 2 e^{\frac{V_{GS(MB)}}{n V_t}}$

¹⁵Die quantitative Berechnung erfolgt in Abschnitt 3.2.4

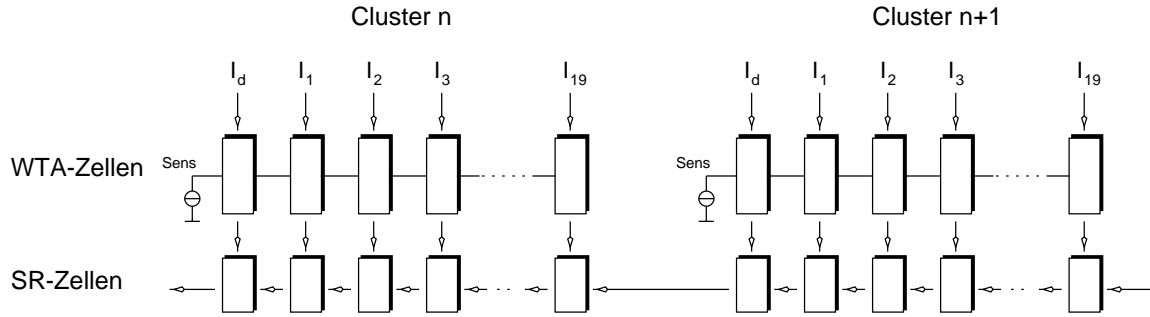


Abbildung 3.19: Die Verbindung von WTA-Zellen und Schieberegister-Zellen

Die hauptsächliche Maßgabe zum Entwurf der Schieberegister-Zellen entstammt der Raumforderung des HSSX. Die Anordnung der WTA-Zellen und der Schieberegister-Zellen erfolgt zwischen den Pixel-Matrizen der Cluster; daher dürfen die Zellen zur Optimierung der optisch aktiven Fläche eines Cluster nur eine minimale Layout-Fläche beanspruchen. Die Schieberegister sollen zum Einen eine Speicher-Zelle zum Speichern der geladenen WTA-Informationen umfassen, zum Anderen eine Daisy-Chain-Architektur ermöglichen, um die Daten an den Kompressor zu übertragen. Der Aufbau der Schieberegister erfolgt daher unter Verwendung einer minimalen Anzahl von MOS-FET durch Realisierung einer *dynamischen* Struktur. In Abbildung 3.20 ist die Anordnung der einzelnen MOS-FET gezeigt. Es handelt sich um ein dynamisches Schiebe-

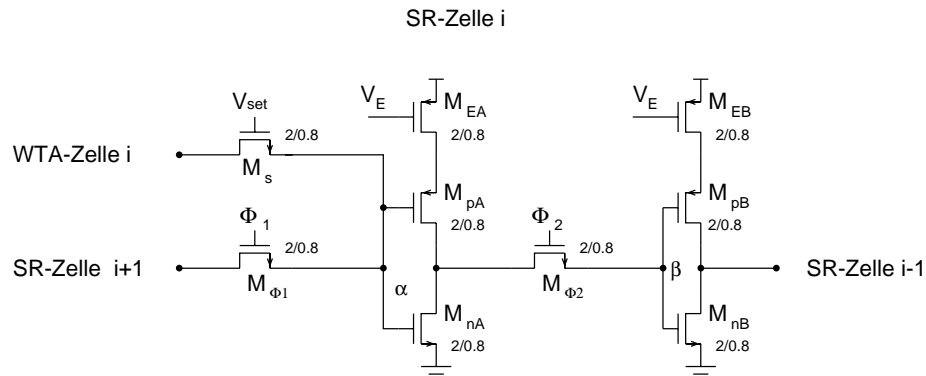


Abbildung 3.20: Der Aufbau einer Schieberegister-Zelle

register mit Zweiphasen-Taktsteuerung und Reset-Funktion und umfasst zwei Inverter A und B, bestehend aus den MOS-FET M_{nA}/M_{pA} und M_{nB}/M_{pB} . Die Inverter werden während des Schiebens der Daten abwechselnd durch ein jeweiliges Aktivieren der Pass-Transistoren $M_{\Phi 1}$ und $M_{\Phi 2}$ angesteuert. Über den Pass-Transistor M_s wird die Ausgangsspannung der zugehörigen WTA-Zelle an den Eingang des ersten Inverters übertragen und ermöglicht ein Laden der Schieberegister-Zelle.

Einführung eines Reset-Transistors

Der Begriff “dynamisch” beschreibt die Flüchtigkeit der Informationen des Schieberegisters bei längeren Wartezyklen. Die Funktionalität ist nur in einem dynamischen Betrieb gesichert. Die Knoten α und β nehmen bei Abschalten einen Zustand hoher Impedanz an, und die Ladung der am Knoten angeschlossenen parasitären Kapazitäten geht über Leckströme der verschiedenen pn-Übergänge verloren. Die Potentiale der Knoten nehmen daher nach einer Weile nicht

definierte Werte an, die unter Umständen die nachfolgenden Inverter in einen stromführenden Betrieb bringen. Der Querstrom, der bei Anlegen eines Eingangs-Spannungswertes im mittleren Spannungsbereich entsteht, beträgt bis zu $210 \mu\text{A}$. Aufgrund der durch die hohe Anzahl der Schieberegister-Zellen im *HSSX* entstehenden Querströme werden Reset-MOS-FET eingeführt, die den Stromfluss der beiden Inverter während der Deaktivierungsdauer abschalten. Abbildung 3.21 verdeutlicht die in einem Inverter entstehenden Querströme bei Anlegen verschiedener Eingangsspannungen. Der Graph wurde durch Simulation eines Inverters mit MOS-FET der Dimensionierung $S = 2/0.8$ unter Verwendung der beiden Worst-Case-Parametersätze *ws* und *wp*¹⁶ des $0.6 \mu\text{m}$ -CMOS-Prozess von *AMS* erstellt. Unter Annahme besonders “stark” ausfallender MOS-

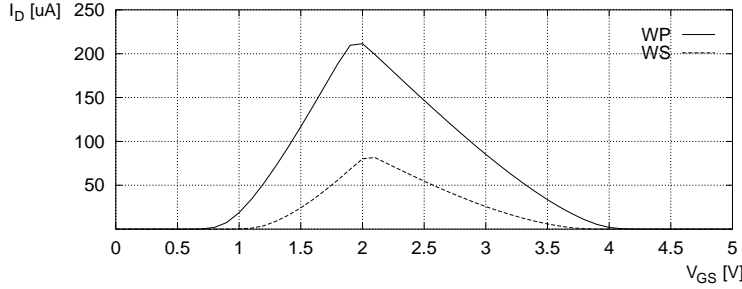


Abbildung 3.21: Der simulierte Querstrom eines Inverters

FET können Querströme bis zu $210 \mu\text{A}$ entstehen, so dass sich unter Summation der Querströme aller 10240 Inverter des *HSSX* ein maximaler Strom von 2.1 A ergeben würde. Die mit diesem Strom verbundene Wärmeleistung würde die Zuleitungen der Spannungsversorgungs-Leitungen der dynamischen Schieberegister zerstören.

Überhöhung der Schaltpotentiale von Φ_1 und Φ_2

Die Funktion der Schieberegister muss im Hinblick auf die Verwendung von Pass-Transistoren und die daraus resultierenden reduzierten Signalpegel der Knoten α und β genauer betrachtet werden. Wie im vorangehenden Abschnitt angesprochen, überträgt ein MOS-FET als Pass-Transistor Eingangsspannungen nur bis zu einem gewissen Spannungspegel. Dies ist eine Konsequenz der Schwellspannung des Transistors. Unter Annahme eines Betriebes des MOS-FET an der Grenze von linearem Bereich und Sättigungsbereich fällt zwischen Gate und Source eine Spannung von

$$V_{GS} = \sqrt{\frac{2 I_D}{\beta}} + V_{THn} \quad (3.46)$$

ab. Die Schwellspannung V_{THn} ist durch den Einfluss des Body-Effektes verändert und beträgt

$$V_{THn} = V_{THn0} + \gamma(\sqrt{V_{BS} + 2 \Phi_f} - \sqrt{2 \Phi_f}). \quad (3.47)$$

Die Variablen β , Φ_f und γ repräsentieren Charakteristika des MOS-FET im Überschwellspannungs-Bereich, siehe [Allen 87c]. Die Spannung V_{GS} reduziert das Potential des Source-Knotens entsprechend, so dass an den Knoten α und β ein maximales Potential von $\Phi_1 - V_{GS(M\Phi_1)}$ und $\Phi_2 - V_{GS(M\Phi_2)}$ anliegen kann. Die Spannung V_{GS} der Pass-Transistoren kann unter entsprechenden Parameter-Schwankungen des Herstellungs-Prozesses bis zu 2 V betragen¹⁷. Weiterhin

¹⁶wp (worst-case-power): starke NMOS-FET, starke PMOS-FET, ws (worst-case-speed): schwache NMOS-FET, schwache PMOS-FET

¹⁷ $\gamma \approx 0.90 \sqrt{V}$, $V_{THn0} \approx 0.95 \text{ V}$, aus [AMS 0.6 I]

kommt es beim Abschalten des Pass-Transistors zu einem Effekt, der als *Clock-Feedthrough* bezeichnet wird. Dieser Effekt beschreibt das kapazitive Koppeln der Gate- und Source-Elektrode des NMOS-FET und reduziert während einer negativen Spannungs-Änderung der Gate-Source-Spannung aufgrund von Ladungsausgleich das Source-Potential um einen weiteren abzuschätzenden Betrag. Nach [Allen 87d] ist dieser Betrag

$$\Delta V_S = -\frac{C_{GS}}{C_{GS} + C_L} (V_D + V_{THn}). \quad (3.48)$$

C_L ist die Summe der an dem Source-Knoten des Pass-Transistors angeschlossenen Kapazitäten und beträgt unter Berücksichtigung der Source-Sperrschicht-Kapazität des Pass-Transistors und der beiden Gate-Kapazitäten des Inverters 11.0 fF. C_{GS} ist die Überlapp-Kapazität zwischen Gate- und Source-Elektrode des Pass-Transistors und berechnet sich zu 0.68 fF. Bei Anlegen eines Drain-Potentials von ca. 5 V erfolgt ein Spannungshub von $\Delta V_s \approx -0.4 V$.

Das Potential des Source-Knotens würde entsprechend der Beträge von V_{THn} , ΔV_s und des Gate-Potentials von 5 V eine Höhe von ca. 2.6 V annehmen können, so dass ein nicht zu vernachlässigender Querstrom innerhalb des Inverters entstehen kann. In Abbildung 3.21 lässt sich bei einer Eingangsspannung von 2.6 V einen Querstrom von ca. 130 μA abschätzen. Der Querstrom tritt innerhalb einer Schieberegister-Zelle nur in einem Inverter auf, da der jeweils andere Inverter die invertierte Eingangsspannung erhält. Diese wird durch den Pass-Transistor in der ursprünglichen Höhe von einigen mV übertragen, so dass der Querstrom des entsprechenden Inverters zu vernachlässigen ist. Die Schieberegister-Zellen einer Cluster-Matrix-Zeile oder -Spalte werden über eine Versorgungsleitung angeschlossen, die den Querstrom der 16:20 Schieberegisterzellen von $I = \frac{320}{2} \cdot 130 \mu A = 20.8 mA$ führen muss. Dies entspricht einem über der Zeit gemittelten Strom von 20.8 μA , da die Schieberegister nur zu 1% der Betriebszeit des HSS aktiv sind.

Nach den Design-Regeln aus [AMS 0.6 II] beträgt die Spitzenbelastungsfähigkeit einer Leiterbahn den 10-fachen Betrag des zeitlich integralen Stromes, so dass die Zuleitung der Inverter einer Cluster-Matrix-Zeile eine Leiterbahnbreite von 2.1 μm bei 1 mA/ μm aufweisen müsste. Diese Breite ist im Rahmen der Raumanforderungen nicht akzeptabel. Eine Abhilfe wäre die Verwendung eines zweiten, parallelen Pass-Transistors in Form eines PMOS-FET mit komplementärem Eingangssignal, um ein Transmission-Gate aufzubauen. Dadurch würden sich die Signalpegel der Knoten α und β über den vollen Versorgungsspannungs-Bereich einstellen lassen und die resultierenden Querströme der Inverter wären vernachlässigbar. Diese Möglichkeit kommt jedoch aufgrund der zusätzlich benötigten Layout-Fläche nicht in Betracht. Der Querstrom der Inverter wird daher durch eine Anhebung der Schaltpotentiale Φ_1 und Φ_2 reduziert. Bei Anhebung der Potentiale um einen Betrag von 0.5 V ergibt sich eine Eingangsspannung der Inverter von minimal 3.1 V, und nach Abbildung 3.21 resultiert ein Spitzenstrom von maximal 75 μA , einem Gesamt-Querstrom der Inverter einer Cluster-Matrix-Zeile oder -Spalte entsprechend von $I = \frac{320}{2} \cdot 75 \mu A = 12 mA$. Eine akzeptable Leiterbahnbreite von 1.2 μm ist erforderlich.

Die Maximale Taktrate der Schieberegister

Nach den Ausführungen des Abschnitts 3.1 ist es notwendig, die Daten der WTA-Zellen, die nach dem Detektionsvorgang in die Schieberegister-Zellen geladen wurden, mit einer Taktrate von 32 MHz zu schieben. Das Einstellen der Potentiale auf stabile Werte muss folglich in einer Zeitspanne von weniger als 31.25 ns stattfinden. Es erfolgt die Bestimmung der maximalen Taktrate der dynamischen Schieberegister. Nach Anlegen eines Potentials an dem Eingang eines Inverters stellt sich die Spannung des nachfolgenden Invertereinganges mit der Verzögerung eines RC-Gliedes ein. Der Widerstand des RC-Gliedes wird durch die Kanalwiderstände der MOS-FET des Treiber-Inverters und des Pass-Transistors gebildet, die Kapazität durch die parasitären

Kapazitäten der Drain-Gebiete des Inverters und des Pass-Transistors und den Gate-Kapazitäten des Last-Inverters. Abbildung 3.22 stellt diese Elemente dar. Für die Berechnung der RC-

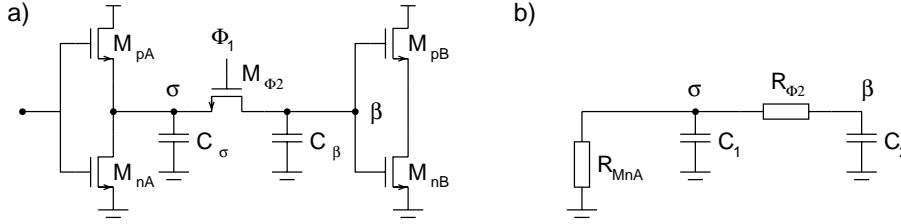


Abbildung 3.22: Berechnung der maximalen Taktrate

Konstante τ sei angenommen, dass der Inverter eine stabile Ausgangsspannung eingestellt hat. Das Potential des Knotens σ betrage 0 V, das des Knotens β betrage 3.1 V. Nach Aktivierung des Pass-Transistors durch Anheben des Potentials Φ_2 auf 5.5 V verringert sich der vorher gegen unendlich gehende Widerstand des Pass-Transistors und die Kapazität C_β entlädt sich über die Widerstände R_{MnA} und R_{Φ_2} . Die maximale Taktgeschwindigkeit wird durch die zeitliche Dauer zum Erreichen eines Potentials von 36% des Endwertes (entsprechend 1.8 V) bestimmt. Nach Gleichung 3.42 werden die Widerstände der Transistoren M_{MnA} und M_{Φ_2} durch $R \approx \beta(V_{GS} - V_{TH})$ angenähert. Unter Verwendung von $\beta = 1.4/0.6 \cdot 100 \mu A/V^2$ ergibt sich für R_{MnA} mit $V_{TH(MnA)} = 0.84 V$ und $V_{GS} = 5 V$ ein Wert von ca. 1030 Ω , für R_{Φ_2} mit $V_{TH(MnA)} = 1.95 V$ und $V_{GS} = 5.5 V$ ein Wert von ca. 1200 Ω . Die Kapazitätswerte lassen sich abschätzen zu

$$\begin{aligned} C_\sigma &= C_{jD(MnA)} + C_{jD(MpA)} + C_{jD(M\Phi_2)} \quad \text{und} \\ C_\beta &= C_{jD(M\Phi_2)} + C_{G(MnB)} + C_{G(MpB)}. \end{aligned} \quad (3.49)$$

Einsetzen der Prozess- und Layout-Parameter ergibt $C_\sigma \approx 8.34 fF$ und $C_\beta \approx 7.63 fF$. Die RC-Konstante beträgt $\tau = 35 ps$. Die Zeitdauer zum Einstellen der Potentiale ist sehr viel kürzer als in der Spezifikation gefordert. Tatsächlich ist der limitierende Faktor der Taktrate des Schieberegisters nicht die interne Funktionalität der Inverter selbst, sondern vielmehr die sog. *Clock-Skew*¹⁸ der Takt-Versorgung von Φ_1 und Φ_2 . Diese Problematik wird in Abschnitt 3.2.5 behandelt.

Das Ende der Schieberegister-Kette einer Cluster-Zeile oder -Spalte wird durch ein Inverterpaar gebildet. Der Ausgang des zweiten Inverters ist zum einen direkt an einen Ausgangstreiber des *HSSX* angeschlossen, um das Signal trotz einer Fehlfunktion der Kompressors durch das externe System weiterverarbeiten zu können. Zum anderen befinden sich die entsprechenden MOS-FET-Gates des Kompressors an dem Ende der Daisy-Chain, verbunden durch eine lange Metallbahn. Das Inverterpaar entkoppelt den Inverter der letzten Schieberegister-Halbzelle von der parasitären Kapazität dieser Elemente. Das Inverterpaar ist aus zwei Standardzellen der 0.6 μm -CMOS-Standardzellen-Bibliothek von AMS aufgebaut und die Dauer $\Delta t_{\Phi_2(dyn)}$ der zeitlichen Verzögerung zwischen der positiven Takt-Flanke von Φ_2 zur Ansteuerung des zweiten Pass-Transistors der Zelle und der Signaländerung der Metallbahn zur Verbindung des Kompressors berechnet sich aus der Standardzellen-Verzögerungszeit unter Anschluss einer definierten Kapazität. Die parasitäre Leiterbahnkapazität der Metallbahn beträgt $C_{pm} = 1.4 \mu m \cdot 7600 \mu m \cdot 0.033 fF/\mu m^2 + 2 \cdot 7600 \mu m \cdot 0.046 fF/\mu m = 1.01 pF$. Die Summe der an der Metallbahn angeschlossenen Gate-Kapazitäten ist 0.16 pF. Der zweite Inverter besitzt nach [AMS 0.6 III] bei Belastung mit $C_L = 1.17 pF$ unter Annahme einer Anstiegsdauer des Eingangssignales von 0.1 ns eine Verzögerungszeit von 1.66 ns. Die Verzögerungsdauer des Inverters der

¹⁸Verzögerung der Taktimpulse aufgrund von unsymmetrischen Belastungen der Clock-Treiber

Schieberegister-Halbzelle und des ersten Inverters nach der Schieberegisterzelle betragen unter Annahme einer Anstiegsdauer des Eingangssignales von 0.1 ns je ca. 0.2 ns. Es resultiert eine Verzögerungsdauer von $\Delta t_{\Phi_2(dyn)} = 2.06 \text{ ns}$. Diese zeitliche Dauer ist für die Berechnung der maximalen Systemfrequenz in Abschnitt 3.2.7 relevant.

3.2.5 Der Takt-Treiberbaum der dynamischen Schieberegister

Die Taktzuführung der dynamischen Schieberegister muss unter Gewährleistung einer maximalen Taktfrequenz von 32 Mhz ermöglicht werden. Um die Taktsignale Φ_1 und Φ_2 gleichförmig über den *HSSX* zu verteilen, wird eine Treiberstruktur entworfen, die die Anzahl von 10240 Pass-Transistoren aller Halbzellen der Schieberegister mit minimaler zeitlicher Dilatation ansteuern kann. Die Forderung eines Taktes von 32 Mhz definiert die zeitliche Dauer einer Taktperiode von 31.25 ns. Unter Berücksichtigung der Gate-Kapazitäten der Pass-Transistoren mit einer Fläche von $W \cdot L = 1.4 \cdot 0.6 \mu\text{m}^2$ ergibt sich eine Gesamtkapazität von $C_c = 10240 W L C_{ox} \approx 23.7 \text{ pF}$. Zur Ansteuerung dieser Last wird eine kaskadierte Treiber-Kette entworfen, deren einzelne Elemente in dem Layout derart angeordnet sind, dass sich für alle Gates die gleiche Clock-Skew ergibt. Abbildung 3.23 stellt die Anordnung der Treiber zur Generation von Φ_1 des *HSSX* dar.

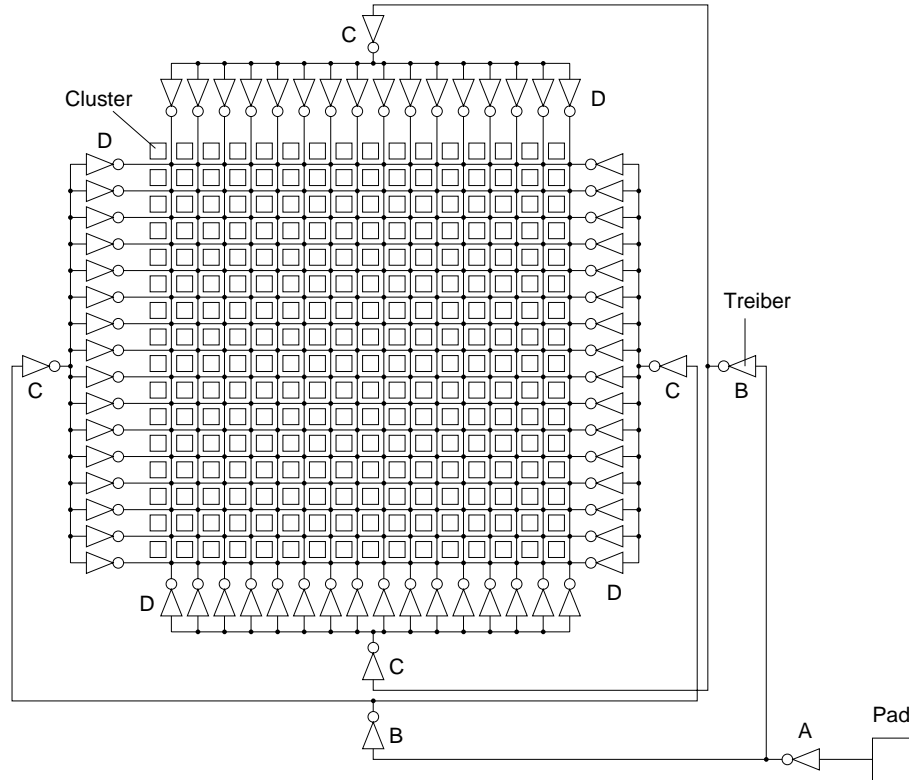


Abbildung 3.23: Der Treiberbaum des Taktes Φ_1

Ausgehend von dem Eingangs-Pad des *HSSX* verzweigt sich der Treiberbaum in 4 Äste. Jeder der 4 Äste steuert 16 Treiber an, die das Taktsignal in die Cluster-Matrix führen. Die Cluster-Matrix wird von allen Seiten gleichzeitig angesteuert. Durch die Verbindung der Taktleitungen an allen Kreuzungspunkten innerhalb der Cluster-Matrix wird gewährleistet, dass die Pass-Transistoren bei annähernd gleicher Eigenschaft aller Treiber zeitlich übereinstimmende Taktflanken erhalten. Die Länge der Signalfade zwischen den Treibern der verschiedenen Klassen

A,B,C und D sind durch die Nachbildung von Kardan-Strukturen identisch. Eine Asymmetrie erfolgt zwischen den Klassen C und D, da die Eingänge der Treiber D über eine gesamte Breite der Cluster-Matrix verteilt sind und sich die Zuleitung des C-Klasse-Treibers in der Mitte der Breite befindet. Innerhalb der Cluster-Matrix entwickelt sich ebenfalls eine Asymmetrie, da die Einspeisung der Taktsignale lediglich an den äußeren Rändern der Matrix erfolgt. Im Folgenden wird die maximale Dilatation der Taktsignale quantifiziert.

Nach [AMS 0.6 I] beträgt die parasitäre Kapazität einer Metall-Leiterbahn mit der Breite W und der Länge L $C_{pm} = W \cdot L \cdot 0.033 \text{ fF}/\mu\text{m}^2 + 2 \cdot L \cdot 0.046 \text{ fF}/\mu\text{m}$. Die parasitären Lastkapazitäten der Metallbahnen der Treiber der verschiedenen Klassen berechnen sich unter Verwendung der Maße $W = 1.4 \mu\text{m}$ und $L = n \cdot 3400 \mu\text{m}$. Die Lastkapazität C_G der Treiber setzt sich aus den Gate-Kapazitäten der angesteuerten MOS-FET zusammen. Die Treiber sind Standard-Inverter der digitalen Standardzellen-Bibliothek des 0.6 μm -CMOS-Prozesses. In [AMS 0.6 III] können die Werte der Eingangskapazitäten erfasst werden. Die beiden Anteile der Lastkapazitäten sind in Tabelle 3.1 für die verschiedenen Treiber-Klassen zusammengefasst. Zur Berechnung des Wertes C_G der Treiber-Klasse D wird eine Anzahl von 160 Pass-Transistoren angenommen, da eine Cluster-Zeile oder -Spalte von 2 Seiten versorgt wird und im Mittel jeweils die Hälfte der Gate-Kapazitäten der insgesamt 320 Pass-Transistoren von einem Treiber versorgt werden.

Klasse	n	C_{pm} [pF]	C_G [pF]	$\sum C$ [pF]	Typ
A	2	1.48	0.53	2.01	IN4
B	4	1.88	0.21	2.09	IN8
C	2	0.94	1.69	2.63	IN4
D	2	0.94	0.38	1.68	IN4

Tabelle 3.1: Lastkapazitäten der Treiber-Klassen. *Typ* bezeichnet die verwendete Inverter-Standardzelle des Prozesses

Die Verzögerungszeit t_d und die Anstiegsdauer der Signal-Flanken t_{sl} der Ausgangssignale der D-Klasse-Treiber bei Definition eines Eingangssignals am Pad lassen sich aufgrund der Tabellenwerte und [AMS 0.6 III] abschätzen und betragen $t_d \approx 4.2 \text{ ns}$ und $t_{sl} \approx 3 \text{ ns}$. Dabei wird eine Anstiegszeit des Eingangssignals am Pad von 2 ns angenommen.

Die Funktionalität der Schieberegister ist gesichert, solange die beiden Taktsignale von Φ_1 und Φ_2 zu keiner Zeit gleichzeitig aktiv sind. Die zeitliche Dauer der Aktivierungs-Phase eines Taktsignales kann maximal $\frac{31.25}{2} \text{ ns} \approx 15.5 \text{ ns}$ betragen. Bei Vorgabe einer Aktivierungs-Dauer von 8 ns und einer Periode von 31.25 ns sind die berechneten Werte von $t_d \approx 4.2 \text{ ns}$ und $t_{sl} \approx 3 \text{ ns}$ ausreichend, um ein zeitliches Überlappen der Aktiv-Phasen zu verhindern.

Die Clock-Skew wird durch die Betrachtung der unsymmetrischen Takt-Zuführungen an die D-Klasse-Treiber innerhalb der Cluster-Matrix abgeschätzt. Die übrigen Treiber unterliegen einer zu vernachlässigenden Clock-Skew, da die Zuführung der Takt-Signale symmetrisch erfolgt. Die Zuleitungen der D-Klasse-Treiber können durch eine RC-Kette nachgebildet werden. Ausgehend von dem Anschlusspunkt des Ausgangs des C-Klasse-Treibers werden daher jeweils 8 RC-Glieder und die jeweils angeschlossenen D-Klasse-Treiber betrachtet. Der Leitungswiderstand der Metallbahn eines RC-Gliedes beträgt unter Annahme einer Leitungslänge von $L = 400 \mu\text{m}$ und $W = 1.4 \mu\text{m}$ nach [AMS 0.6 II] ca. 114 Ω , die Kapazität der Leitung und des angeschlossenen D-Klasse-Treibers ca. 162 fF. Die zeitliche Verzögerung der Glieder untereinander wird zu $\tau \approx 18 \text{ ps}$ abgeschätzt. Die direkt am Ausgang des C-Klasse-Treibers angeordneten D-Klasse-Treiber erhalten das Takt-Signal daher ca. 144 ps vor denen am äußeren Ende des Astes. Diese

Zeitdauer beträgt ca. 5% von t_{sl} und kann daher vernachlässigt werden. Die Clock-Skew innerhalb der Cluster-Matrix ist unter den gleichen Parametern der RC-Glieder zu berechnen und beträgt ebenfalls ca. 5% von t_{sl} . Zwischen allen Schieberegistern entsteht folglich eine maximale Clock-Skew von 5% der Anstiegsdauer der Taktsignale. Diese Clock-Skew ist vernachlässigbar.

Die Verlust-Leistung, die während des Schiebens in den Treibern des HSSX umgesetzt wird, kann durch Abschätzen des integralen Querstromes der CMOS-Inverter, aus denen die Treiber bestehen, berechnet werden. In [AMS 0.6 III] werden die Verlustleistungen der Standardzellen normiert auf die Taktfrequenz definiert. Durch Addition der Werte für die 134 IN_4 - und die 4 IN_8 -Treiber¹⁹ ergibt sich eine Gesamt-Verlustleistung von $P_{tot} \approx 49.4 + 0.74 \text{ mW}$ mit einem integralen Querstrom von ca. 10 mA. Bezogen auf die Wiederholrate von 1 ms ergibt sich während des Betriebs des HSSX eine integrale Strombelastung von 0.1 mA und eine Verlustleistung von 500 μW .

3.2.6 Das Layout der WTA- und Schieberegister-Zelle

Die Abbildung 3.24 stellt den kompletten Schaltplan einer WTA-Zelle mit der zugehörigen Schieberegister-Zelle dar. Es fehlt die Stromquelle I_{src} , die den *Flow*-Strom der M_{Fn} generiert. Der entsprechende MOS-FET, der die Stromquelle implementiert, ist seitlich der X- und Y-Ausleseschaltkreise angeordnet. Die Dimensionierungen aller MOS-FET sind seitlich der jeweiligen Akronyme der MOS-FET angezeigt.

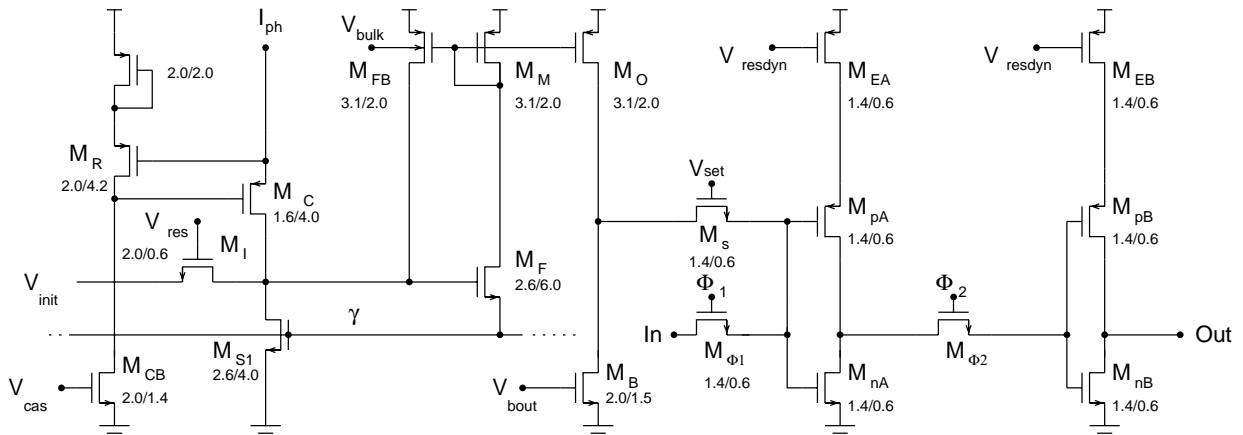


Abbildung 3.24: Das Schematic einer WTA- und SR-Zelle

Abbildung 3.25 zeigt das Layout einer WTA-SR-Zelle. Die Zelle ist in vertikaler Richtung 20-fach kaskadiert und seitlich der Photodetektor-Matrix eines Cluster angeordnet. An der linken Seite ist der Anschluss zur Einspeisung des Photostromes des Photodetektor-Streifens zu sehen. Im Folgenden wird die Anordnung der Elemente des Layouts begründet und die Vorgehensweise zur Reduzierung des Mismatchings der einzelnen MOS-FET erläutert. Die Layout-Fläche der WTA-SR-Zelle beträgt $17.6 \times 63.4 \mu\text{m}^2$. Die Fläche ist in zwei Bereiche unterteilt. Die linken zwei Drittel werden durch die Elemente der WTA-Zelle belegt, das rechte Drittel von den Elementen der SR-Zelle.

Von der linken Seite beginnend sind drei Signalpfade vorhanden, die im Schaltungsbetrieb empfindlich auf parasitäres Übersprechen reagieren. Sie führen die Signale V_{src} , V_{dmy} , und V_{cas} . In

¹⁹Diese Anzahlen umfassen die Treiber der zwei Treiber-Bäume von Φ_1 und Φ_2

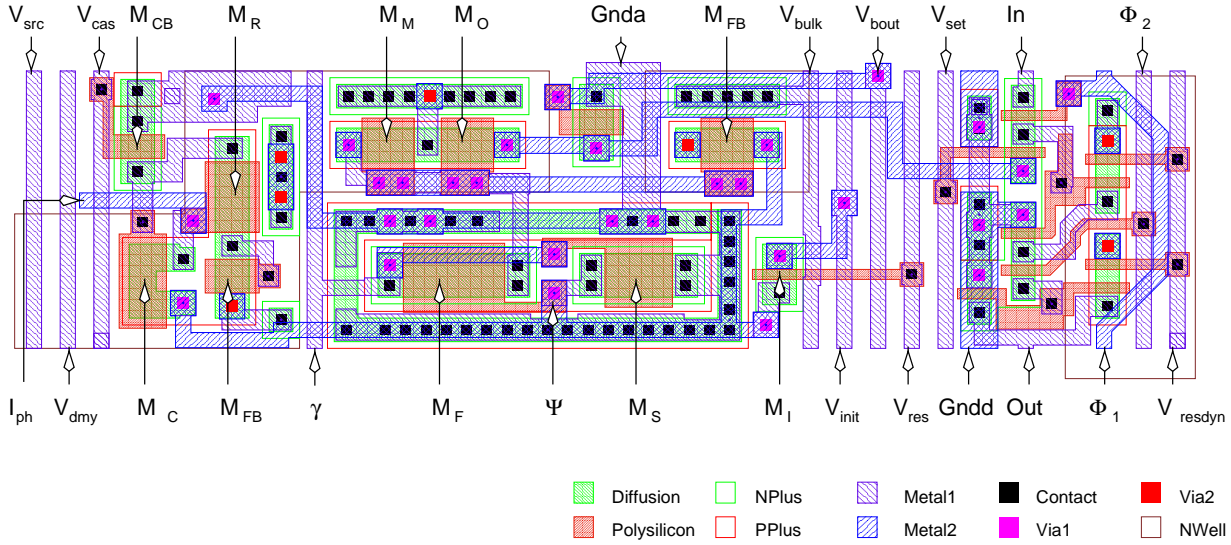


Abbildung 3.25: Das Layout einer WTA- und SR-Zelle. Der Anschluss an den entsprechenden Photodetektor-Streifen der Zelle erfolgt an der linken Seite.

dem Bereich zwischen WTA-Zell-Kern und Photodetektor-Matrix ist das kapazitive Einkoppeln von Potentialänderungen der verschiedenen Signale der WTA-Zelle am geringsten. V_{src} führt die Bias-Spannung der Stromquelle I_{src} , die den Ausgangsstrom der M_{Fn} definiert. V_{dmy} liefert die Bias-Spannung der Dummy-Zelle eines WTA-Moduls, durch die ein definiertes Verhalten des WTA-Moduls bei undefinierter Bestrahlung der Photodetektoren-Matrix erzeugt wird. V_{cas} stellt über den M_{CB} den Ruhestrom der regulierten Kaskode ein. Es folgen die MOS-FET der regulierten Kaskode, die durch den Signalpfad γ von den MOS-FET des Kernes der WTA-Zelle getrennt sind.

Die beiden Transistoren M_F und M_S stellen den Kern der WTA-Zelle dar und befinden sich in einem durch einen geschlossenen Substratkontakt-Ring geschützten Bereich, der das Übersprechen der benachbarten Transistoren reduziert und ein stabiles Substratpotential definiert. Die Fläche des M_F und des M_S ist maximal groß gewählt, da inhärentes Mismatching von MOS-FETs durch Vergrößerung der Fläche reduziert wird. Die Dimensionierung von W/L erfolgt mit $W/L < 1$, um einen geringen Kanallängen-Modulationsfaktor der Transistoren zu erzielen.

Über dem M_F und dem M_S sind die Stromspiegel-Transistoren M_M , M_O und M_{FB} angeordnet, sowie der Bias-Transistor M_B , der den Ruhestrom des Ausgangsbuffers der WTA-Zelle generiert. Der Transistor M_{FB} befindet sich in einer eigenen N-Wanne, da das Bulk-Potential zum Einstellen des Untersetzungsverhältnisses der Feedback-Ströme mit einem von V_{dd} verschiedenen Potential angesteuert wird. Es ist wünschenswert, diesen Transistor in näherer Umgebung des M_M anzuordnen, um Mismatching durch Prozess-Gradienten zu reduzieren. Die Design-Regeln des CMOS-Prozesses nach [AMS 0.6 II] fordern jedoch einen großen Abstand von N-Wannen verschiedener Potentiale, da sonst unerwünschte Effekte eines parasitären npn-Bipolar-Transistors²⁰ auftreten könnten.

Neben M_S befindet sich der Reset-Transistor zur Definition des Start-Potentials des Knotens Ψ . Der Transistor verursacht aufgrund der negativen Spannungs-Flanke des Signales V_{res} ein unerwünschtes Übersprechen auf die sensiblen Schaltungselemente der WTA-Zelle und ist daher

²⁰Dieser Transistor wird durch zwei benachbarte N-Wannen und das Substrat gebildet

so weit wie möglich neben den übrigen Transistoren angeordnet. Daneben folgen die Signalfade V_{bulk} , V_{init} , V_{bout} und V_{res} .

V_{res} ist ein Signalfad, der aufgrund des schnellen Potentialsprunges durch kapazitives Koppeln auf benachbarte Pfade übersprechen kann. Der Pfad ist am weitesten von den übrigen Elementen der WTA-Zelle entfernt. Der Pfad V_{bout} definiert den Ruhestrom des Ausgangsbuffers der WTA-Zelle und ist während des Schaltvorganges von V_{set} nicht aktiv, sondern kurze Zeit später und kann daher direkt neben V_{set} platziert werden. Der Pfad V_{bulk} ist am weitesten von V_{set} entfernt, um eine stabile Bulk-Spannung zu gewährleisten. Im weiteren Verlauf des Layouts befinden sich die Signalfade und die MOS-FET der SR-Zelle. Die Transistoren sind mit $W/L = 1.4/0.6$ dimensioniert, um unter den gegebenen Design-Regeln eine minimale Fläche der MOS-FET zu belegen.

Die Masse-Zuleitungen von WTA-Zelle ($Gnda$) und SR-Zelle ($Gndd$) sind getrennt, die positive Versorgungsspannung ist für beide Zellteile identisch. Die umfassende Trennung von digitalen und analogen Versorgungsleitungen ist aufgrund der zeitlichen Abfolge von Detektions- und Schiebevorgang nicht notwendig. Die Zuführung der Versorgungsspannung V_{dd} geschieht durch die dritte Metallisierungsebene des Prozesses und ist in Abbildung 3.25 nicht dargestellt. Sie wird über die Elemente *Via2* des Layouts angeschlossen. Die Metallschicht ist im *HSSX* über der gesamten Fläche der Cluster-Matrix mit Ausnahme der Photodetektor-Matrizen der einzelnen Cluster definiert, um ein unerwünschtes Bestrahlen von MOS-FETs zu verhindern. Nach [Teodori 97] besitzt eine Metallschicht mit einer Dicke von $1\text{ }\mu\text{m}$ einen Dämpfungsfaktor von ca. $6.6 \cdot 10^3$. Die restlichen transmittierenden Photonen sind in der Anzahl zu gering, um unerwünschtes Verhalten der MOS-FET des *HSSX* zu verursachen.

3.2.7 Der Kompressor

Die 256 Cluster des *HSSX* erzeugen Bitstrings mit einer Länge von jeweils 20 Bits. Die Information über die X und Y-Position des Fokalkpunktes ist in der Stelle der logischen 1 innerhalb der Bitstrings kodiert. Sollte keine auswertbare Fokalkpunkt-Intensität vorliegen, ist das Bit der ersten Stelle logisch 1, da die Dummy-Zelle des WTA-Moduls den stärksten Eingangsstrom detektiert. Im Allgemeinen erzeugt genau eine Zelle des WTA-Moduls eine logische 1 und die übrigen Zellen eine logische 0. In den Fällen, in denen zwei Eingangsströme mit gleicher Stärke vorliegen, ist das Detektieren eines Unterschiedes nicht möglich, so dass beide entsprechenden WTA-Zellen jeweils eine logische 1 erzeugen.

Zwei typische Konfigurationen des Bitstring eines WTA-Moduls sind in Abbildung 3.26 gezeigt. D bezeichnet das Bit der Dummy-Zelle, 1 bis 19 die Bits der Positionsdetektion.

1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
D	1	2																	19
Position																			

Abbildung 3.26: Zwei Bitstrings eines WTA-Moduls

Die Daten sind in der Form eines Bitstring für die externe Signalverarbeitung ungünstig angeordnet, so dass eine Konversion erfolgt. Der *Kompressor* erzeugt aus einem Bitstring ein Datenbyte mit 8 Bits, dass die komplette oben definierte Information des Bitstring in einer verdichteten Form enthält.

Die Funktionalität der Kompression basiert auf der Annahme, dass bei einer Bestrahlung der Cluster-Matrix mit Fokuspunkten durch die WTA-Module eine Position detektiert wird, die entweder durch die Aktivierung genau eines Positions-Bit, oder bei einer Unsicherheit durch maximal zwei Positions-Bits markiert wird. Bei Dominanz der Dummy-Zelle aufgrund einer zu schwachen Intensität des Fokuspunktes wird das Dark-Bit aktiviert und die übrigen Bits des Bitstring deaktiviert. Das Aktivieren von mehr als drei Bits ist als Fehler der Positionsdetektion zu interpretieren. Sollte keines der Bits des WTA-Bitstring aktiviert werden, liegt eine zu schwache Bestrahlung der Photodetektor-Matrix und ein zu schwacher Dummy-Strom vor.

Die Position des Fokuspunktes wird in 19 diskrete Abschnitte unterteilt, so dass die Darstellung der Information unter Binärkodierung durch 5 Positions-Bits erfolgen kann. Die Aktivierung des ersten der 19 Bits wird durch Positions-String 00001₂ dargestellt. Die Aktivierung des Dark-Bits wird als Positions-String 00000₂ definiert. Bei einer Aktivierung von gleichzeitig 2 Bits innerhalb des Bitstring des WTA-Moduls wird ein "Double"-Bit gesetzt und die Position des zweiten der beiden Bits durch die Positions-Bits markiert. Das Auftreten von mehr als 2 Bits wird durch die Aktivierung eines "Error"-Bit angezeigt. Sollte keines der Bits des WTA-Bitstrings aktiviert werden, wird das "Dark"-Bit gesetzt. Die Anzahl der benötigten Bits zur Darstellung des Bitstrings eines WTA-Moduls ist daher in einem Byte kodierbar. Abbildung 3.27 stellt fünf WTA-Bitstrings unter Auswahl der vorher definierten möglichen Konfigurationen dar.

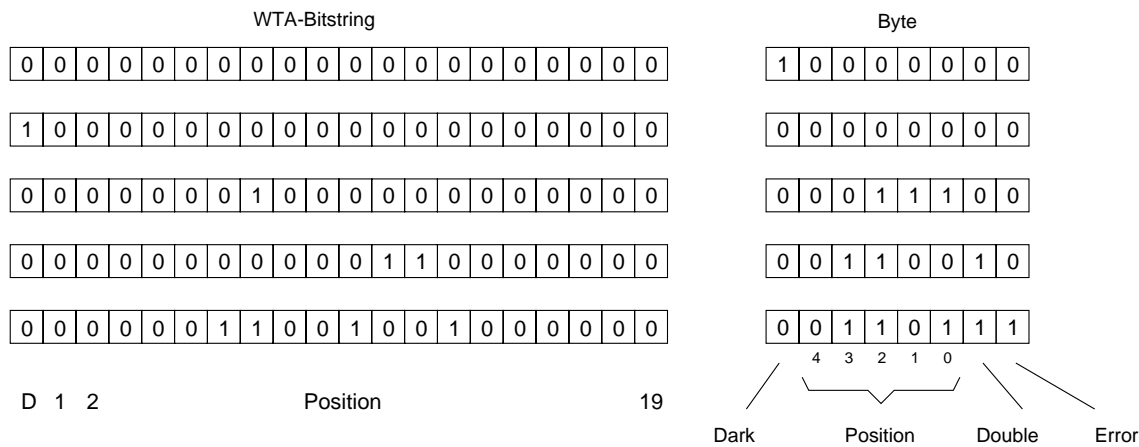


Abbildung 3.27: Schematik der Kompression

Aufgrund der Daisy-Chain-Architektur der Cluster und der Anordnung der Cluster in Zeilen und Spalten mit jeweils 16 Cluster werden beim Schieben der WTA-Bitstrings insgesamt 32 Datenströme erzeugt, die je 16 aufeinanderfolgende WTA-Bitstrings enthalten. Um die Datenmenge in der Zeit während des Schiebens verarbeiten zu können, wird der Kompressor in einer zweistufigen Pipeline-Architektur aufgebaut. Die 32 Datenströme werden parallel verarbeitet, indem zeitgleich in der ersten Stufe die WTA-Bitstrings der 32 Datenströme komprimiert und in der zweiten Stufe die resultierenden Bytes an die Ausgänge des *HSSX* gegeben werden. Dazu werden in einem Durchgang innerhalb von 20 Taktzyklen die sequentiell eintreffenden 20 Bits der 32 Datenströme komprimiert. Der Kompressor ist in zwei Hälften unterteilt, um die Daten für X- und Y gleichzeitig handhaben zu können. Die generierten 32 Bytes (16 X-Daten und 16 Y-Daten) der ersten Stufe werden vor Beginn des nächsten Durchganges an die 32 Schieberegister der zweiten Stufe übergeben und während der ersten 16 Taktzyklen des nächsten Durchganges an die Ausgänge-Ports des *HSSX* geschoben. Es erfolgen insgesamt 16 Durchgänge.

Um diese Struktur des Kompressors zu realisieren, wird ein Kompressions-Modul entwickelt, das die 20 Bits eines WTA-Bitstring sequentiell verarbeiten und nach dem Eintreffen des letzten der 20 Bits das Ergebnis zur Verfügung stellen kann. Die vorangehend vorgestellte Funktion der Kompression muss dazu durch eine logische Verknüpfung der schon verarbeiteten Bits mit dem aktuellen Bit des Bitstring implementiert werden, so dass das Ergebnis vor dem Eintreffen des ersten Bit des folgenden WTA-Bitstring feststeht.

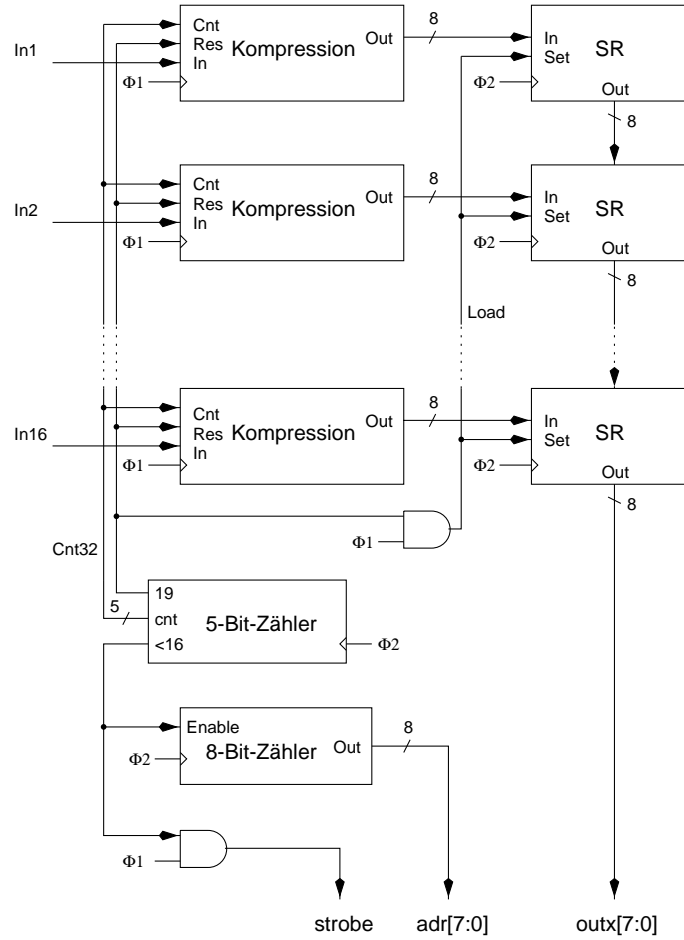
Das Kompressions-Modul verwendet zur Implementierung der vorangehend beschriebenen Algorithmik einen synchronen 5-Bit-Zähler und ein synchrones logisches Verknüpfungsnetzwerk. Der Zähler ist als 5-Bit-Zähler mit asynchronem und synchronem Reset realisiert. Er wird nach einem Reset ausgehend von einem Zählerstand 0 bei einer positiven Taktflanke von Φ_2 um 1 inkrementiert und nach Erreichen des Zählerstandes 10011_2 (19) beim Eintreffen der nächsten positiven Taktflanke auf 0 gesetzt.

Das Ergebnis der Kompression wird in einem 8-bit-Register *Out* gespeichert, das zu Beginn eines Durchganges das Byte 10000000_2 enthält. *Out*[7] stellt das Dark-Bit dar, *Out*[6..2] die Position des Fokalkpunktes, *Out*[1] das Double-Bit und *Out*[0] das Error-Bit. Der Datenstrom wird am Eingang des Kompressions-Moduls abgetastet. Bei Auftreten einer logischen 1 am Eingang wird das Register *Out*[7] (Dark-Bit) gelöscht und der Zählerstand in die Register *Out*[6:2] übertragen. Das Register *Out*[1] (Double-Bit) wird mit dem Eingangssignal und dem Register *Out*[7] logisch verknüpft. Beim Eintreffen einer logischen 1 wird der Stand von *Out*[0] überprüft und bei einem Inhalt von logisch 0 wird *Out*[1] gesetzt. *Out*[0] (Error-Bit) wird mit *Out*[1] logisch verknüpft. Beim Eintreffen einer logischen 1 am Eingang wird *Out*[1] überprüft und bei einem Inhalt von logisch 1 wird *Out*[0] gesetzt. Die Register *Out*[1] und *Out*[0] werden bei Eintreten des Zählerstandes 0 zur Initialisierung des nächsten Durchganges gelöscht, *Out*[7] gesetzt. Der Stand der Positionsregister *Out*[6:2] wird erhalten, um gegebenenfalls beim Vorkommen einer zu geringen Fokalkpunktintensität des nächsten Cluster einen möglichen Wert der Interpolation der Position zu geben.

Der Zähler erzeugt nach Erreichen des Zählerstandes 10011_2 bei einer positiven Taktflanke von Φ_1 ein Set-Signal, mit dem die Werte der Register *Out*[7:0] in ein Schieberegister geladen werden. Durch Triggern des Signals auf die Taktflanke von Φ_1 wird sichergestellt, dass die Daten, die erst kurze Zeit nach Eintritt der vorangehenden Taktflanke von Φ_2 zur Verfügung stehen, schon bei Beginn der nächsten Taktflanke von Φ_2 durch die Schieberegister geschoben werden können, während die Werte des neuen Durchganges bereits erzeugt werden.

Das Schieben der Daten wird durch einen vollsynchronen 8-Bit-Zähler mit asynchronem Reset gesteuert, der nach dem ersten Durchgang der Kompression gestartet wird. Beim Eintreffen einer positiven Flanke von Φ_2 wird der Zähler inkrementiert, solange der Zählerstand des 5-Bit-Zählers kleiner 16 ist. Dadurch wird ein Zählen der tatsächlich am Ausgang der Schieberegister der Kompressionseinheit gültigen Bytes ermöglicht. Nach Erreichen des Zählerstandes 256 wird das Signal *Ready* erzeugt, das dem externen System einen vollständigen Durchgang der Kompression meldet. Der Zählerstand wird über die Signale *Adr*[7:0] an das äußere System übertragen, ebenso wie ein Signal *Strobe*, das durch ein logisches AND des Taktsignales Φ_1 und des Signales, das den 5-Bit-Zählerstand 16 anzeigt, gebildet wird. Die beiden Signale *Adr*[7:0] und *Strobe* werden verwendet, um die Kompressions-Daten *OutX*[7:0] und *OutY*[7:0] ohne äußere Steuerung in ein RAM des externen Systems übertragen zu können (siehe Kapitel 4).

Die Abbildung 3.28 zeigt den schematischen Aufbau der Hälfte des Kompressors, die zur Kompression der X-Positions-Daten verwendet wird. Die Eingänge In_1 bis In_{16} werden mit den X-Ausgängen der Inverterpaare verbunden, die das jeweilige Ende einer Daisy-Chain der Cluster-Matrix darstellen.

Abbildung 3.28: Der Kompressor des *HSSX*

Der Kompressor ist vollsynchron aufgebaut, so dass alle Signale ihre Zustände nur auf eine positive Flanke der Taktsignale Φ_1 oder Φ_2 ändern. Eine Ausnahme bildet lediglich das Signal *Strobe*, das intern keine Funktion besitzt. Die einwandfreie Funktionalität der digitalen Schaltung ist gewährleistet, solange keine Race-Condition²¹ erfüllt wird. Im Folgenden wird die Clock-Skew des Kompressors quantifiziert.

Zur Zuführung der Taktsignale Φ_1 und Φ_2 wird ein Treiberbaum nach Abbildung 3.29 entwickelt, der das zeitsynchrone Triggern der im Kompressor enthaltenen Flip-Flops ermöglicht und das Eintreten einer Race-Condition verhindert. In einem rückgekoppelten synchronen logischen Schaltwerk liegt die Hauptanforderung der Signalpfad-Definition in der symmetrischen Zuführung des Takt-Signales an die getriggerten Elemente wie Flip-Flops oder Latches. Die zeitgleiche Übernahme der Eingangs-Signale verhindert Fehlschaltungen aufgrund von unterschiedlichen Laufzeiten zwischen Triggersignal und Logiksignal. Das Schaltwerk des Kompressors ist derart aufgebaut, dass die maximale Dilatation einer Taktflanke in der gesamten Architektur des Kompressors kleiner ist als die minimale Zeitverzögerung in der Signalverarbeitung der verwendeten Flip-Flops.

Tabelle 3.2 stellt die verwendeten Treiber und ihre Lastkapazitäten unterteilt in Metallbahn-Kapazität und Gate-Kapazitäten der angeschlossenen Logik-Bausteine dar. Die Metallbahn-

²¹ Das Eingangssignal eines Netzwerkes ändert seinen Zustand innerhalb eines Taktzyklus, bevor das Taktsignal seinen Zustand ändert.

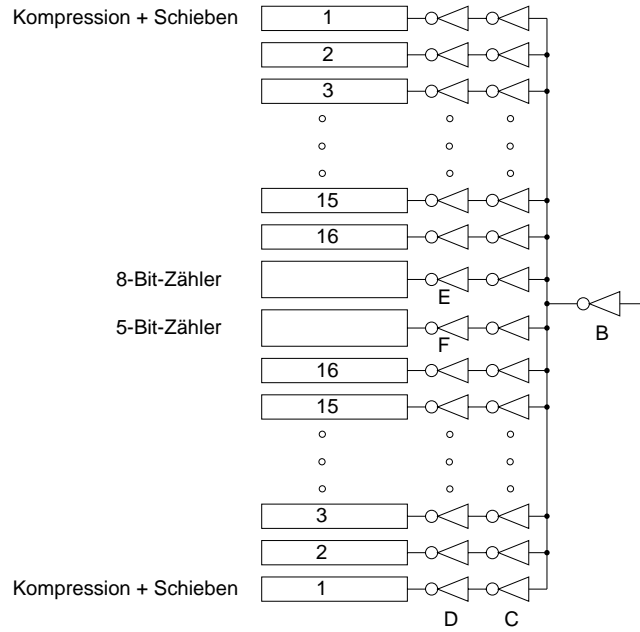


Abbildung 3.29: Der Treiberbaum des Taktes Φ_2 des Kompressors. Die Abbildung stellt die jeweils 16 Kompressions-Module der beiden Detektions-Richtungen und die beiden Zähler dar.

Kapazitäten der Treiber der Klassen C, D, E und F betragen weniger als 2 fF und werden daher vernachlässigt. Die Verzögerungszeit und die Anstiegsdauer der Signal-Flanken der Ausgangs-

Klasse	C_{pm} [pF]	C_G [pF]	$\sum C$ [pF]	Typ
A	1.48	0.53	2.01	IN4
B	0.45	0.99	1.43	IN4
C	-	0.05	0.05	IN1
D	-	0.18	0.18	IN2
E	-	0.19	0.19	IN2
F	-	0.10	0.10	IN2

Tabelle 3.2: Lastkapazitäten der Treiber-Klassen des Kompressors von Takt Φ_2

signale der D-, E- und F-Klasse-Treiber bei Definition eines Eingangssignals am Pad lassen sich unter Berücksichtigung der Verzögerungszeiten der A-, B- und C-Klasse-Treibern, der Tabellenwerte und [AMS 0.6 III] abschätzen und betragen $t_{dD} \approx 3.35 \text{ ns}$, $t_{dE} \approx 3.35 \text{ ns}$ und $t_{dF} \approx 3.20 \text{ ns}$ bei jeweils $t_{slD} = t_{slE} = t_{slF} = 1.05 \text{ ns}$. Die maximale zeitliche Dilatation des Taktsignales beträgt zwischen den Ausgängen der D-, E- und F-Klasse-Treibern 0.15 ns. Dabei wird eine Anstiegszeit des Eingangssignals am Pad von 2 ns angenommen. Im Kompressor werden Flip-Flops des Types *DFSA*, *DF8* und *DFA* verwendet. Die minimale Signalverzögerungszeit zwischen der positiven Flanke des Taktsignales und des Ausgangssignales dieser Elemente beträgt nach [AMS 0.6 III] für ein *DF8* 1.0 ns, für ein *DFA* 1.3 ns und für ein *DFSA* 1.12 ns bei Belastung des Ausganges mit $C_L = 0.1 \text{ pF}$. Für alle Flip-Flops des Kompressors ist daher die Bedingung erfüllt, dass die maximale Zeitdilatation der Takt-Flanke des Trigger-Signals kleiner als die minimale Zeitverzögerung der Flip-Flops ist. Eine Race-Condition liegt unter keinen Umständen

vor.

Die maximale Taktfrequenz des Kompressors bestimmt sich aus der Verzögerungsdauer des längsten Signalpfades innerhalb des Kompressors und der zeitlichen Charakteristik der Eingangssignale, die den Kompressor mit der Cluster-Matrix verbinden. Unter allen Modulen des Kompressors verfügt der 5-Bit-Zähler über den längsten Signalpfad, dessen Verzögerungszeit unter Addition aller im Pfad befindlichen Elemente ca. 2.5 nsec beträgt. Unter Berücksichtigung der Anstiegszeit des Signales von ca. 2.0 ns und der Setup-Zeit des Flip-Flops der Zählerzelle von ca. 0.4 ns, die zu der Verzögerungszeit des Signalpfades addiert wird, ergibt sich eine maximale Zeitdauer der Signalverzögerung nach der Taktflanke von Φ_2 von ca. 3.9 ns. Der Zähler könnte mit einer Taktfrequenz von ca. 220 MHz betrieben werden.

Zur Erläuterung der zeitlichen Abfolge der Signale des Kompressions-Moduls sei Abbildung 3.30 betrachtet. Das Kompressions-Modul benötigt als Eingangssignale den WTA-Datenstrom einer Cluster-Zeile- oder Spalte, den Zählerstand des 5-Bit-Zählers und drei Signale, die von den eigenen Ausgangssignalen zurückgeführt werden. Die Übernahme der Eingangssignale zur Generation der Ausgangssignale darf erst nach einer Stabilisierung der Eingangssignale erfolgen. Die zeitliche Dauer der Stabilisierung wird im Folgenden abgeschätzt, um die maximale Systemfrequenz zu berechnen. Die Zeitdauer t_1 beschreibt den zeitlichen Abstand zwi-

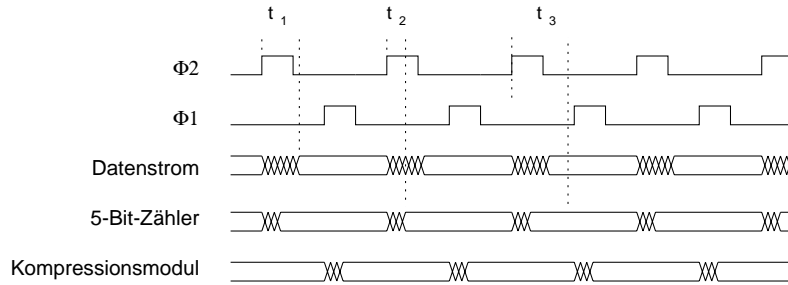


Abbildung 3.30: Zeitdiagramm des Kompressors

schen positiver Taktflanke von Φ_2 und einer Stabilisierung des Signales des Datenstromes. Der Ausgang des dynamischen Schieberegisters der letzten Zelle einer Daisy-Chain steuert ein Inverter-Paar, um die parasitäre Kapazität der Zuleitung des Kompressors in einer kurzen Zeitdauer umladen zu können. Unter Betrachtung der in Abschnitt 3.2.4 berechneten Dauer $\Delta t_{\Phi_2(dyn)} = 2.06 \text{ ns}$ und der Verzögerungszeit des Treiberbaumes der dynamischen Schieberegister lässt sich $t_1 = 4.2 \text{ ns} + 2.06 \text{ ns} \approx 6.3 \text{ ns}$ abschätzen.

Die Zeitdauer t_2 beschreibt den zeitlichen Abstand zwischen positiver Taktflanke von Φ_2 und einer Stabilisierung der Signale des 5-Bit-Zählers. Jede Zählerzelle verwendet ein flankengetriggertes Flip-Flop des Typs *DFA* und zwei nachfolgende Inverter *IN2* und *IN4*, um die 32 Kompressions-Module ansteuern zu können. t_2 lässt sich unter Berücksichtigung der Verzögerungszeit des Flip-Flops, der Verzögerungszeit von *IN2* und der Verzögerungszeit von *IN4* unter Annahme einer Lastkapazität²² von 1.2 pF berechnen. Nach [AMS 0.6 III] errechnet sich eine Verzögerungszeit der Elemente von ca. 3 ns, so dass sich zusammen mit der Verzögerungszeit des Kompressor-Treiberbaumes $t_2 = 6.3 \text{ ns}$ ergibt.

Die Zeitdauer t_3 beschreibt die sog. Setup-Zeitdauer, in der die Eingangssignale einen stabilen Zustand besitzen müssen, um fehlerfrei in die Flip-Flops der Kompressions-Module übertragen zu werden. Diese Zeitdauer beträgt bei einem Flip-Flop des Typs *DFSA* ca. 0.4 ns. Unter

²² Dies schließt die parasitäre Leiterbahn-Kapazitäten und die Eingangskapazitäten der 32 Kompressions-Module ein.

Betrachtung der Verzögerungszeit des Treiberbaumes des Kompressors ergibt sich eine Zeitdauer von ca. 2.95 ns. Aus diesen Zeiten lässt sich der minimale Abstand der beiden Taktflanken von Φ_2 und Φ_1 ableiten, der zu einer Funktionsfähigkeit des Kompressions-Moduls notwendig ist. Der zeitliche Abstand Δt_Φ der positiven Flanken von Φ_2 und Φ_1 muss mindestens

$$\Delta t_\Phi > \max(6.3 \text{ ns}, 6.35 \text{ ns}) - 2.95 \text{ ns} = 3.4 \text{ ns} \quad (3.50)$$

betragen, um eine Stabilität der Eingangssignale des Kompressions-Moduls zu gewährleisten. Die Funktionalität des Kompressors wäre bei einem symmetrischen Zweiphasentakt mit einer Frequenz von $f = \frac{1}{2 \cdot \Delta t_\Phi} = 147 \text{ MHz}$ möglich. Bei Verwendung des Systemtaktes von 32 MHz ist die Funktionalität des Kompressors als gesichert anzusehen.

Die Verlustleistung des Kompressors während der Kompression berechnet sich aus der Summe der Verlustleistung aller Standardzellen. Unter Verwendung von [AMS 0.6 III] ergibt sich eine frequenzbezogene Verlustleistung von $803.78 \mu\text{W}/\text{MHz}$, so dass sich bei einer Systemfrequenz von 32 MHz eine Verlustleistung von 25.7 mW bei einem Strom von 5.14 mA ergibt. Die Leistung ist bezogen auf einen Dauerbetrieb des Kompressors. Die Verlustleistung während des Betriebes des HSSX beträgt $257 \mu\text{W}$, die integrale Strombelastung beträgt $51.4 \mu\text{A}$.

3.2.8 Die Steuerung des HSSX

Aus den Erläuterungen der letzten Abschnitte lässt sich die zeitliche Abfolge der Steuersignale des HSSX zum Erzeugen einer Bildinformation ableiten. Abbildung 3.31 stellt das Zeitdiagramm der wichtigsten logischen Signale dar. Der Zeitpunkt t_3 sei das Ende der Detektionsphase einer Positionsbestimmung der Fokulpunkte, so dass die Information der WTA-Zellen bereitsteht und in die Schieberegister geladen werden kann. Der Schiebe- und Kompressionszyklus beginnt bei

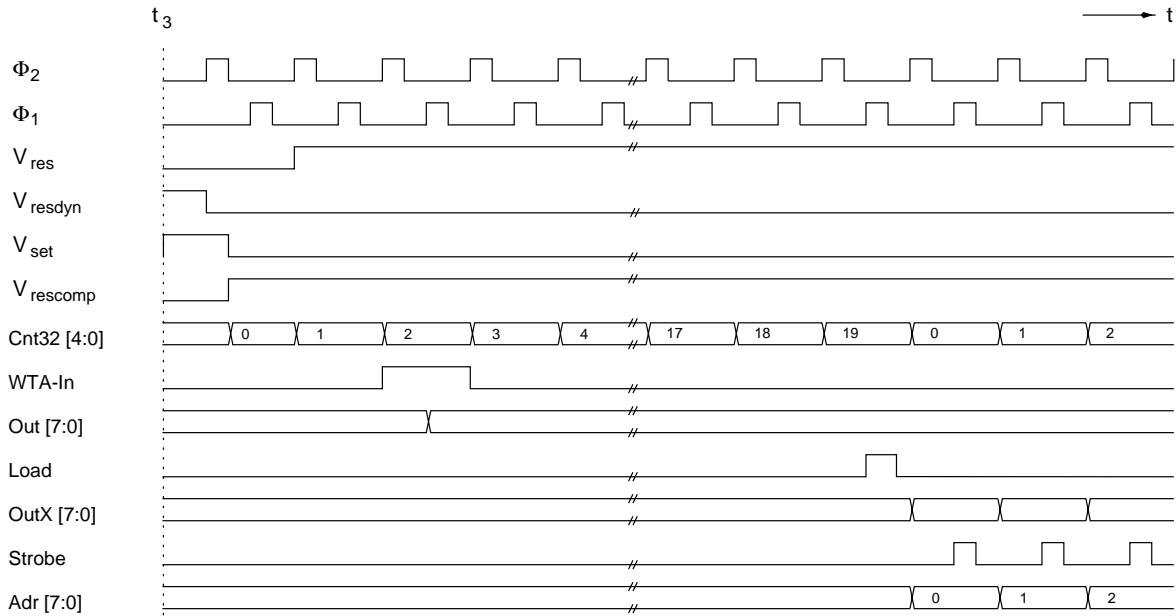


Abbildung 3.31: Zeitdiagramm des HSSX

Eintritt der ersten positiven Taktflanke von Φ_2 mit der Deaktivierung des Reset-Signales V_{resdyn} der Schieberegister-Inverter. Eine definierte Zeitdauer von ca. $10 \mu\text{s}$ zuvor wird das Signal V_{set} zum Laden der WTA-Informationen in die Schieberegister aktiviert. V_{set} ist bis zum Ende

der ersten aktiven Phase von Φ_2 logisch 1, um die Information durch die beiden geöffneten Pass-Transistoren M_s und $M_{\Phi 2}$ an den Ausgang der entsprechenden Schieberegister-Zelle zu übertragen. Zum Ende der ersten aktiven Phase von Φ_2 wird V_{rescomp} deaktiviert, um den Kompressor seine Funktion aufnehmen zu lassen. Bei Eintreffen der ersten positiven Taktflanke von Φ_1 kann der Zählerstand 0 des 32-Bit-Zählers unter möglicher Aktivierung des Bit 1 des WTA-Bitstring in die Register des Kompressions-Moduls geladen werden. Beim Eintreffen der nächsten positiven Taktflanke von Φ_2 muss das Signal V_{res} deaktiviert werden, um die WTA-Zellen die Initialisierung der Potentiale vornehmen zu lassen.

Es schließen sich die 19 Taktzyklen an, um den entsprechenden Ausgangswert $\text{Out}[7:0]$ des Kompressions-Moduls zu erzeugen. Nach Erreichen des Zählerstandes 19 wird an einer positiven Taktflanke von Φ_1 das Signal Load aktiviert, um den Wert von $\text{Out}[7:0]$ in die entsprechende Schieberegister-Zelle des Kompressors zu laden. Das Signal steht beim Eintreffen der nächsten positiven Taktflanke von Φ_2 an den Ausgängen $\text{OutX}[7:0]$ zur Verfügung. Gleichzeitig beginnt der 8-Bit-Zähler mit der Erzeugung der Adresse in $\text{Adr}[7:0]$ und das Strobe-Signal bestätigt die Gültigkeit der beiden Daten-Bytes an den Ausgängen $\text{OutX}[7:0]$ und $\text{OutY}[7:0]$. Nach Ablauf der gesamten Kompression werden die Signale V_{res} , V_{resdyn} und V_{rescomp} wieder deaktiviert und es erfolgt eine erneute Detektion der Fokalkpunkt-Positionen.

3.2.9 Die Ein- und Ausgabe-Pads des HSSX

Der *HSSX* wird über Ein- und Ausgangs-Pads der 0.6 μm -CMOS-Standardzellen-Bibliothek mit dem externen System verbunden. Die Cluster-Matrix verfügt über Eingangspads für die 5 analogen Spannungen V_{src} , V_{dmy} , V_{cas} , V_{bulk} , V_{init} und V_{bout} , und die 5 digitalen Signale $V_{\Phi 1}$, $V_{\Phi 2}$, V_{set} , V_{res} und V_{resdyn} . Zu Testzwecken können die Eingangs-Zellen der dynamischen Schieberegister über digitale Eingangstreiber angesteuert werden und die Datenströme der Daisy-Chains über Ausgangstreiber ausgelesen werden. Die entsprechenden Ein- und Ausgangstreiber werden über dedizierte Versorgungsleitungen angeschlossen, um die erhöhte Verlustleistung der Test-Pads im normalen Betrieb des *HSSX* durch Abschalten der Treiber zu vermeiden. Der Kompressor verfügt über einen zusätzlichen Eingangs-Pad für das digitale Signal V_{rescomp} . Die Signale $V_{\Phi 1}$ und $V_{\Phi 2}$ werden gemeinsam mit der Cluster-Matrix verwendet. Die Signale $\text{Adr}[7:0]$, Strobe, $\text{OutX}[7:0]$ und $\text{OutY}[7:0]$ des Kompressors werden über die digitalen Ausgangs-Pads an das externe System übertragen. Es handelt sich um Ausgangstreiber des Typs *OB33*, die einen Laststrom von 4 mA treiben können.

Bei Betrieb des Kompressors wird in den Ausgangstreibern eine Verlustleistung umgesetzt, die nach [AMS 0.6 III] 292 $\mu\text{W}/\text{MHz}$ beträgt. Unter Annahme einer Taktfrequenz von 32 MHz ergibt sich im Dauerbetrieb der 26 Ausgangstreiber des Kompressors unter Ansteuerung von CMOS-kompatiblen Lasten eine Verlustleistung von 243 mW bei einem Strom von 48.5 mA. Bei Betrieb des *HSSX* beträgt die Verlustleistung im zeitlichen Mittel 2.4 mW mit einem integralen Strom von 0.48 mA. Ein Kurzschluss der Ausgangstreiber aller 26 Signalleitungen würde einen Dauerstrom von maximal 104 mA verursachen, da jeder der Treiber bis zu 4 mA führen kann. Die Versorgungsleitungen der Ausgangstreiber sind mit 60 μm Breite in der Lage, diesen Strom ohne Zerstörung zu führen.

Die Ausgangstreiber der Daisy-Chain-Datenströme werden für die X- und die Y-Daten über getrennte Versorgungsleitungen angeschlossen. Es handelt sich um jeweils 16 Ausgangstreiber, so dass sich eine Dauerleistung von 299 mW und ein Dauerstrom von 59 mA bei einer Taktfrequenz von 32 MHz errechnet. Bei Betrieb des *HSSX* beträgt die Verlustleistung im zeitlichen

Mittel 3.0 mW mit einem integralen Strom von 0.6 mA. Die Breite der Versorgungsspannungszuleitungen ist ebenfalls ausreichend dimensioniert, um einen Kurzschluss-Strom aller 32 Ausgangstreiber ohne Zerstörung zu führen.

Alle Eingangs-Pads der analogen und digitalen Spannungen sind der Standardzellen-Bibliothek des 0.6 μm -CMOS-Prozesses entnommen und verfügen über Schutzstrukturen gegen ESD²³-Einflüsse. Es handelt sich um die Typen *IB33C* und *IOA3C*. Die digitalen Eingangs-Pads verwenden CMOS-kompatible Pull-Up-Widerstände, so dass bei fehlendem Eingangssignal ein definierter Zustand des logischen Signals vorliegt.

Die Versorgungsspannungen werden über Power-Pads der Standardzellen-Bibliothek des 0.6 μm -CMOS-Prozesses angeschlossen, die ebenfalls über Schutzstrukturen gegen ESD verfügen. Der *HSSX* ist vollständig geschützt gegen ESD-Spannungen von 1000 V nach dem Standard MIL-STD-883C (aus [AMS 0.6 IV]).

3.2.10 Das Layout des HSSX

Abbildung 3.32 zeigt das Layout eines *HSSX*. Die größte Fläche des Chips wird von der Cluster-Matrix abgedeckt. Die 16×16 Cluster sind bis auf die Flächen der Photodetektoren zum Abschirmen der hochsensitiven Schaltungselemente der WTA-Module vollständig mit einer Metall-Lage bedeckt. Am rechten Rand des Chips ist die Struktur des Kompressors zu sehen. Darüber und darunter befinden sich Teststrukturen, um Transistor-Parameter messen zu können (siehe Kapitel 6).

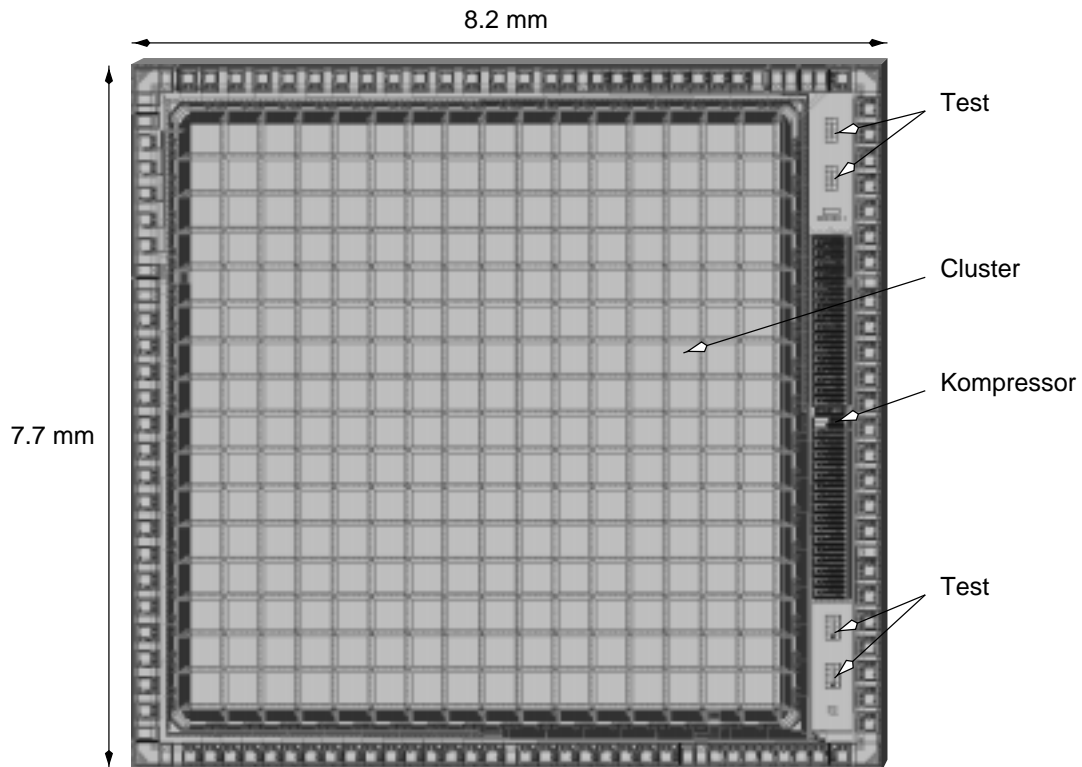


Abbildung 3.32: Der *HSSX*

²³Electro Static Discharging

3.3 Der Prototyp 1

Die folgenden beiden Abschnitte beschreiben qualitativ die beiden Teststudien *Prototyp 1* und *Prototyp 2*, die in der ersten Phase des Projektes entworfen wurden. In den beiden ASICs wurde das Konzept der Fokalkpunkt-Positions-Detektion evaluiert, das auf der Anwendung eines WTA-Modul basiert, wobei in *Prototyp 2* eine erweiterte Signalverarbeitung zur Verbesserung der lateralen Auflösung der Fokalkpunkt-Position eingesetzt wird.

Abbildung 3.33 zeigt eine Fotografie des *Prototyp 1*. In diesem Abschnitt wird die Architektur des *Prototyp 1* vorgestellt.

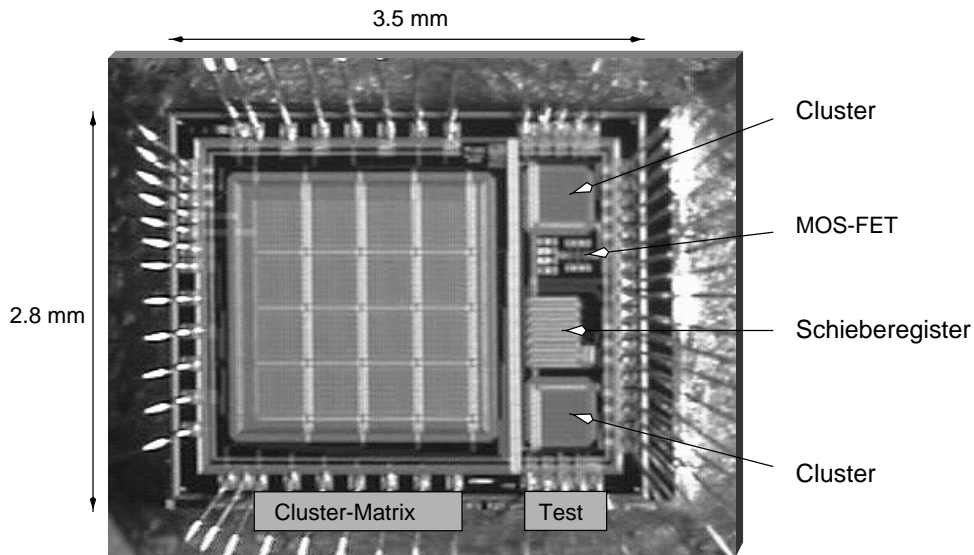


Abbildung 3.33: *Prototyp 1*

Prototyp 1 wurde in der $0.8\ \mu\text{m}$ -CMOS-Technologie der Firma AMS prozessiert. Der größere linke Bereich zeigt die Cluster-Matrix, der kleinere rechte Teil einige Test-Strukturen. Der prinzipielle Aufbau des Chips ist bereits in Abbildung 3.1 von Abschnitt 3.1 dargestellt. Die Kompressionseinheit ist aufgrund der geringen Datenmenge nicht implementiert.

Der Chip besitzt eine Cluster-Matrix von 4×4 Cluster. Jedes Cluster ist in 8×8 Photodetektor-Paare unterteilt, so dass sich eine Anzahl von 8 Photodetektor-Streifen in X- und Y-Richtung ergibt. Die Photodetektoren bestehen aus n^+ /sub-Dioden. Die Kantenlänge eines Cluster beträgt $400 \times 400\ \mu\text{m}^2$, die Breite eines Photodetektor-Streifens $45\ \mu\text{m}$. Die optisch aktive Fläche eines Cluster ist im Vergleich zu dem HSSX etwas größer, da die zwischen den Clustern angeordnete Signalverarbeitung aufgrund der einfacheren Schaltungsstruktur weniger Fläche benötigt.

Zur Signalverarbeitung der Photoströme der Photodetektor-Streifen wird ein WTA-Modul verwendet, dass ohne Rückkopplung und Initialisierung arbeitet. Die generierten Daten der WTA-Module werden über statische Schieberegister an das externe System übertragen. Aufgrund der geringeren Auflösung der Positions-Detektion mit 8 Diskretisierungsstufen und der geringeren Anzahl der Cluster werden die Daten in ihrer ursprünglichen Form übertragen und nicht komprimiert. Der Chip arbeitet mit einer Auflösung von $45\ \mu\text{m}$ bei einer Mindest-Intensität eines Fokalkpunktes von $50\ \text{nW}$ mit einer Wiederholrate von $1\ \text{kHz}$. Im Folgenden wird das WTA-Modul und das statische Schieberegister des Chips erläutert.

Das WTA-Modul des Prototyp 1

Das im *Prototyp 1* implementierte WTA-Modul basiert auf der in Abbildung 3.10 dargestellten WTA-Grundstruktur unter Entkopplung der Photodetektor-Sperrschichtkapazität durch eine Kaskode-Konfiguration. Abbildung 3.34 stellt das im *Prototyp 1* verwendete WTA-Modul vor.

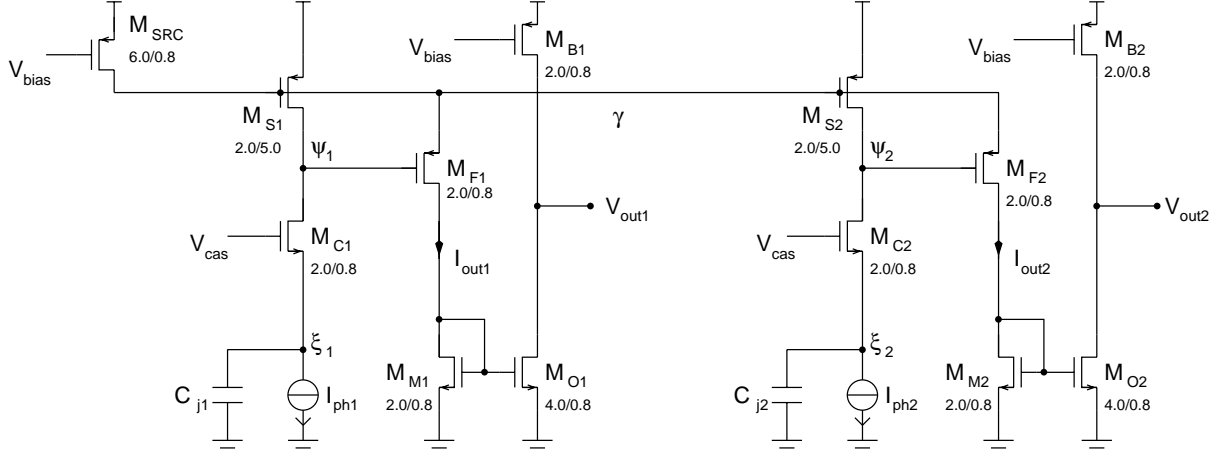


Abbildung 3.34: Das WTA-Modul des *Prototyp 1*

Die Kaskode ist durch einen einzelnen Transistor realisiert, so dass die Entkopplung der Potentialänderung der Knoten ξ_n von den Potentialänderungen der Knoten Ψ_n gegenüber der regulierten Kaskode des *HSSX* um den Faktor $\lambda n V_t$ weniger effizient ist.

Die Spannung $V_{GS(MC1)}$ bestimmt das Potential des Knotens Ψ_1 . Unter Änderung der Drain-Source-Spannung des M_{C1} resultiert eine Änderung der Gate-Source-Spannung von

$$\frac{dv(\xi_1)}{dv(\Psi_1)} = -\frac{dv_{gs(MC1)}}{dv(\Psi_1)} = -\lambda n V_t. \quad (3.51)$$

Die gesamten parasitären Kapazitäten des Knotens Ψ_1 summieren sich zu

$$C_{\Psi_1} = \lambda n V_t C_{jn} + C_{jD(MC1)} + C_{jD(MS1)} + C_{G(MF1)} \quad (3.52)$$

Unter Verwendung von [AMS 0.8] ergibt sich bei einer Fläche eines Photodetektor-Streifens von $8 \cdot 14.8 \cdot 36.5 \mu m^2$ und einer Sperrspannung von 2 V eine Sperrschicht-Kapazität $C_{j1} = 2.18 pF$. Einsetzen der Parameter λ , n und V_t von $0.3 V^{-1}$, 1.5 und 0.025 V ergibt eine für den Knoten Ψ_1 scheinbare Kapazität von 24.5 fF. Addition von $C_{jD(MS1)} = 3.63 fF$, $C_{jD(MC1)} = 2.25 fF$ und $C_{G(MF1)} = 3.45 fF$ ergibt $C_{\Psi_1} = 33.8 fF$.

Die Antwortzeit einer WTA-Zelle auf eine Stromänderung der Eingangsströme von $I_{ph1} - I_{ph2} = \Delta I_{ph}$ berechnet sich analog zu den Ausführungen in Abschnitt 3.2.3. Unter Voraussetzung eines statischen Betriebs der Schaltung mit $I_{ph2} > I_{ph1}$ ändert sich bei Eintritt der Bedingung $I_{ph2} < I_{ph1}$ das Drain-Potential von M_{S1} aufgrund der Photostrom-Differenz ΔI_{ph} . Die Zeitverzögerung zur Änderung des Drain-Potentials um 1 V beträgt bei $C \approx 33.8 fF$ und $\Delta I_{ph} \approx 1 pA$ ca. $\tau \approx 33.8 ms$. Bei Eintritt der Bedingung $V_{DS(MS1)} = V_{DS(MS2)}$ nach ca. 33 ms vertauschen sich die Beträge der Ausgangsströme I_{out1} und I_{out2} und das Ausgangssignal kann abgerufen werden.

Die WTA-Zelle besitzt zwei analoge Eingangsspannungen V_{bias} und V_{cas} . V_{cas} definiert das Gate-Potential der M_{Cn} . Es wird im Betrieb des *Prototyp 1* auf eine Spannung von ca. 2.5 V

gelegt, so dass durch den Spannungsabfall der Gate-Source-Spannungen der M_{Cn} eine Sperrspannung der Photodetektoren von ca. 2 V resultiert. Über V_{bias} wird zum einen der Strom I_{ref} eingestellt, der als Summe der Ausgangsströme I_{outn} fließen kann, zum anderen der Bias-Strom des Ausgangs-Buffers M_{On} und M_{Bn} . Durch die Wahl der Parameter $S_{M(M)}/S_{M(O)} = 2$ und $S_{M(SRC)}/S_{M(B)} = 3$ ergibt sich unter Annahme von $\lambda = 0.3$, $n = 1.5$ und $V_t = 0.025 V$ für beide MOS-FET des Buffers eine Ausgangsspannung einer “verlierenden” WTA-Zelle von $V_{out} \rightarrow 5 V$. Ein “gewinnende” Zelle erzeugt eine Ausgangsspannung von $V_{out} \approx 13 mV$. Das Ausgangssignal einer WTA-Zelle ist low-aktiv. V_{bias} wird im Betrieb des *Prototyp 1* auf ca. 4.40 V gelegt, um einen Strom I_{src} von ca. 50 nA zu definieren.

Die Informationen der WTA-Module werden nach einer Detektionsphase an das statische Schieberegister übertragen. Das Schieberegister verfügt über einen Eingangs-Multiplexer, der wahlweise die Ausgangsspannung der zugehörigen WTA-Zelle oder des vorangehenden Schieberegisters auf den Eingangsknoten des ersten Inverters führen kann. Abbildung 3.35 stellt die Anordnung der MOS-FET des Schieberegisters dar.

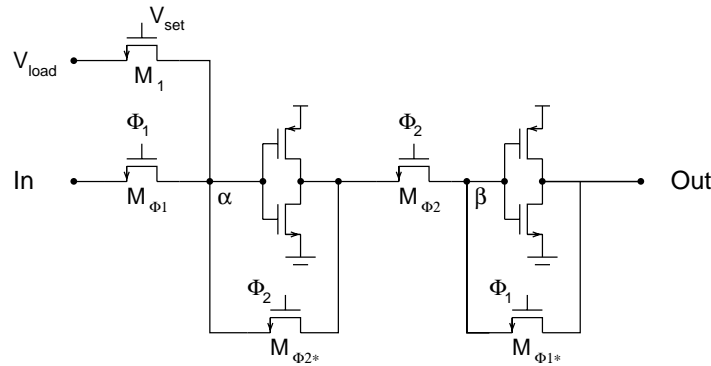


Abbildung 3.35: Das statische Schieberegister des *Prototyp 1*

Die Ansteuerung zum Laden oder Schieben der Informationen geschieht über Pass-Transistoren. Dazu wird ein nichtüberlappender Zweiphasentakt verwendet, wie in Abschnitt 3.2.4 beschrieben. Das Eingangs-Potential des Inverters der ersten Halbzelle wird durch Aktivieren von V_{set} oder Φ_1 über die Pass-Transistoren $M_{\Phi1}$ oder M_1 definiert. Gleichzeitig wird über den Pass-Transistor $M_{\Phi1*}$ der Ausgang des zweiten Inverters an dessen Eingang zurückgeführt, um ein Entladen des Eingangsknotens β zu verhindern. Das System kann in diesem Zustand angehalten werden, da die Potentiale beider Knoten α und β statisch definiert sind. Um die Information des ersten Inverters an den zweiten zu übertragen, wird Φ_2 aktiviert. Dadurch überträgt der Pass-Transistor $M_{\Phi2}$ das Ausgangssignal des ersten Inverters an den Knoten β . Gleichzeitig wird über den Pass-Transistor $M_{\Phi2*}$ der Ausgang des ersten Inverters an dessen Eingang zurückgeführt, um ein Entladen des Eingangs-Knotens α zu verhindern. Der Betrieb eines statischen Schieberegisters ermöglicht einen variablen Systemtakt mit $0 Hz < f_t < f_{max}$. Über die jeweilige Definition der Knotenpotentiale werden die in Abschnitt 3.2.4 besprochenen langandauernden hohen Querströme bei undefiniertem Eingangspotential der Inverter verhindert.

Prototyp 1 wurde im Rahmen des Projektes erfolgreich als Sensor für einen Augenbewegungsverfolger eingesetzt. In Kapitel 5 erfolgt eine Zusammenfassung der Messergebnisse.

3.4 Der Prototyp 2

Die Erkenntnisse aus der Entwicklung des *Prototyp 1* in Bezug auf optische Sensitivität und Auflösung der Verschiebungs-Detektion der Fokalfunkte des Linsenarray wurden verwendet, um einen zweiten Prototyp des Wellenfrontsensor-ASIC zu entwickeln. Die in *Prototyp 1* erreichbare Auflösung in der Bestimmung der Fokalfunkte-Position beträgt $45\ \mu\text{m}$ bei einem dynamischen Bereich von $360\ \mu\text{m}$, die minimal erforderliche Photonenleistung eines Fokalfunktes zum Erreichen der Auslesefrequenz von $1\ \text{kHz}$ beträgt ca. $50\ \text{nW}$. Der Fokus der Entwicklung von *Prototyp 2* lag in der Verfeinerung der Genauigkeit der Positions-Detektion auf möglichst unter $20\ \mu\text{m}$, so dass eine Rekonstruktion der Wellenfront mit einer Genauigkeit von besser als $\pm 0.05\ \text{dpt}$ möglich ist. Gleichzeitig war es von Interesse, die optische Sensitivität zu erhöhen, um die in den Spezifikationen von Kapitel 1.6 geforderte Intensität eines Fokalfunktes von $J_{FP} \approx 200\ \text{pW}/\text{Spot}$ bei gleicher Aufnahme Frequenz verarbeiten zu können. Abbildung 3.36 zeigt eine Fotografie des Chips. In diesem Abschnitt wird die Architektur des *Prototyp 2* vorgestellt.

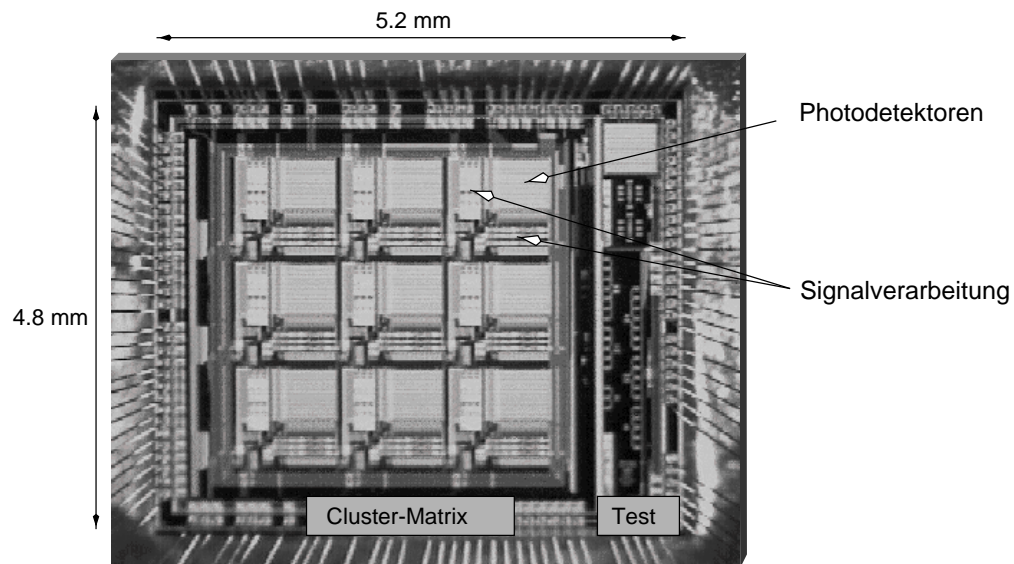


Abbildung 3.36: *Prototyp 2*

Prototyp 2 ist in der $0.8\ \mu\text{m}$ -BiCMOS-Technologie der Firma AMS prozessiert. Die BiCMOS-Technologie stellt entsprechend den Ausführungen von Kapitel 2 einen vertikalen npn-Bipolartransistor zur Verfügung, der als Phototransistor gegenüber den verschiedenen Photodioden eine erhöhte effektive spektrale Empfindlichkeit besitzt. Der linke Bereich des Chip enthält die Cluster-Matrix, im rechten Teil sind einige Teststrukturen angeordnet. Der prinzipielle Aufbau des Chips ist in Abbildung 3.1 von Abschnitt 3.1 dargestellt. Die Kompressor-Einheit ist im *Prototyp 2* nicht implementiert, da die Datenmenge eine Kompression nicht erforderlich machte.

Der Chip besitzt eine Cluster-Matrix von 3×3 Clustern. Jedes Cluster ist in 10×10 Photodetektor-Paare unterteilt, so dass sich eine Anzahl von jeweils 10 Photodetektor-Streifen in X- und Y-Richtung ergibt. Die Breite eines Photodetektor-Streifens beträgt $50\ \mu\text{m}$. Die Kantenlänge eines Cluster beträgt $1061 \times 1061\ \mu\text{m}^2$, da eine größere Fläche für die im Vergleich zum *Prototyp 1* komplexere Singnalverarbeitung verwendet wird. Die optisch aktive Fläche ist daher im Vergleich zum *Prototyp 1* kleiner.

Die Signalverarbeitung des *Prototyp 2* verwendet zur Erhöhung der Genauigkeit der Positionsbestimmung der Fokalfpunkte eine Analog-Digital-Wandlung von Photoströmen, bei gleichzeitiger Detektion des stärksten Photostromes der Photodetektor-Matrix durch ein WTA-Modul. Die digitalisierte Information über die Höhe des Photostromes der “gewinnenden” Zelle, sowie über die Höhen der Photoströme der zwei zur “gewinnenden” Zelle benachbarten Zellen werden über Schieberegister aus dem Chip geschoben, um extern weiterverarbeitet zu werden.

3.4.1 Auflösungsverbesserung durch Interpolation

Die umfangreichere Signalverarbeitung des *Prototyp 2* zielt auf die Möglichkeit zur Interpolation der Fokalfpunkte-Positionen ab. Die Form der Intensitätsverteilung ähnelt einer Airy-Funktion, wie durch die Gleichung 1.12 in Abschnitt 1.6.2 angenähert. Aufgrund der Kenntniss von mehreren Funktionswerten und dem theoretischen Verlauf der Intensität als Funktion der Abszissen-Koordinate kann auf die Lage des Intensitäts-Maximums geschlossen werden. Dieser Vorgang wird im Allgemeinen als Interpolation bezeichnet und gibt die Möglichkeit zur Erhöhung der Auflösung über die nominale, durch die Abstände der Detektions-Streifen definierte Auflösung der Positionsdetektion. Die Güte der Interpolation hängt dabei sehr stark von der tatsächlichen Form der Intensitätsverteilung ab.

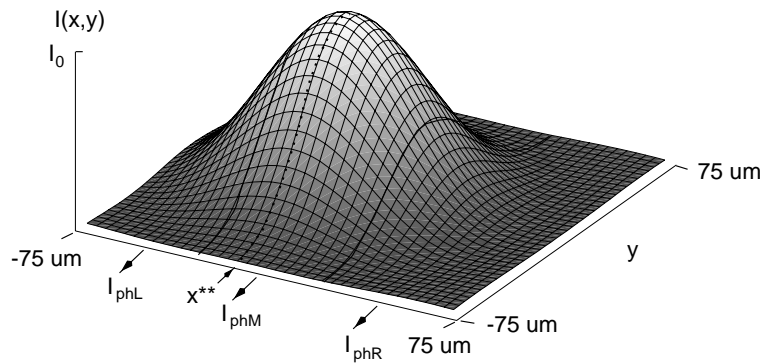


Abbildung 3.37: Interpolation der Positionsdetektion

Abbildung 3.37 stellt den Verlauf der Intensitätsverteilung eines Fokalfpunktes dar. Je nach tatsächlicher Koordinate des Intensitätsmaximums x^{**} ergibt sich nach einer Integration der drei das Zentrum der Intensitätsverteilung umschließenden Photodetektor-Streifen eine charakteristische Verteilung der drei Photoströme. Das WTA-Modul detektiert den mittleren Streifen.

Die Daten der Tabelle 3.3 wurden durch numerische Integration von Gleichung 1.15 unter Wahl von 5 Werten von x^{**} berechnet. x^{**} bezieht sich dabei eine Nullposition, die genau auf der Grenze eines Photodetektor-Streifens zum Nachbarstreifen liegt. Da die Breite eines Streifens $50 \mu m$ beträgt, werden die Koordinaten $x^{**} \in \{5, 15, 25, 35, 45\} \mu m$ gewählt. I_{phL} , I_{phM} und I_{phR} geben die Anteile der Ströme des linken, des mittleren und des rechten Streifens bezogen auf die Summe der drei Ströme wieder.

Durch die Signalverarbeitung des *Prototyp 2* werden die drei Stromhöhen I_{phL} , I_{phM} und I_{phR} mit einer nominalen Auflösung von 10 Bit digitalisiert und zusammen mit der Information x^{*} des WTA-Moduls über die Lage des Intensitätsmaximums an das externe System übergeben.

x^{**} [μm]	5	15	25	35	45
I_{phL}	40.3%	23.8%	11.5%	04.5%	01.7%
I_{phM}	58.0%	71.7%	77.0%	71.7%	58.0%
I_{phR}	01.7%	04.5%	11.5%	23.8%	40.3%

Tabelle 3.3: Verteilung der Ströme

Das externe System kann über den Vergleich der drei Werte mit Einträgen in einer Lookup-Table²⁴ auf die tatsächliche Koordinate x^{**} schließen und zusammen mit dem Wert von x^* des WTA-Moduls die Position des Fokalkpunktes berechnen.

3.4.2 Die Signalverarbeitung des Cluster

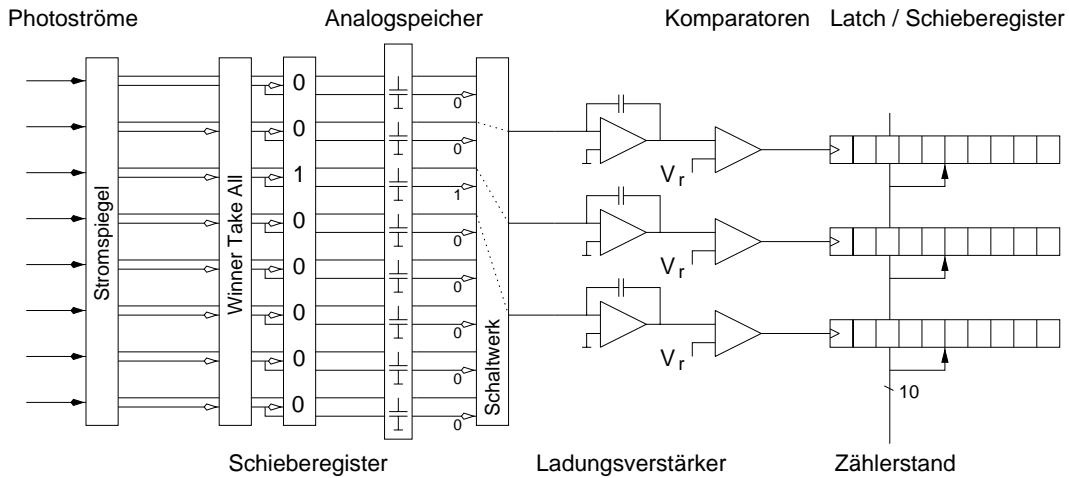
Abbildung 3.38: Signalverarbeitung des *Prototyp 2*

Abbildung 3.38 zeigt die Anordnung des Signalverarbeitungselemente²⁵ eines Cluster des *Prototyp 2*. Die Photoströme der Photodetektor-Streifen werden je durch einen Stromspiegel dupliziert. Der eine Ausgangsstrom des Spiegels wird an die zugehörige WTA-Zelle geleitet, der andere zum Aufladen eines Kondensators verwendet. Nach Ende der Detektion wird das Ausgangssignal der "gewinnenden" WTA-Zelle verwendet, um durch ein Schaltwerk die Ladungsmengen des eigenen und der beiden benachbarten Kondensatoren über drei Ladungsverstärker in Spannungen zu konvertieren. Die Spannungen werden nach diesem Vorgang mit einer definiert steigenden Referenzspannung verglichen und über ein Single-Slope-Verfahren in digitale Werte umgewandelt. Nach Vollenden der AD-Wandlung werden die drei 10 Bit-Werte und die Information des WTA-Moduls durch eine Daisy-Chain an die Ausgänge des ASIC geschoben. Die Schieberegister verwenden die Architektur von *Prototyp 2*.

²⁴Die Funktionswerte als Funktion eines Argumentes werden in einer Tabelle abgelegt. Unter Vergleich des gemessenen Wertes mit den Einträgen der Tabelle kann das Argument gefunden werden.

²⁵In der Abbildung sind aus Gründen der Übersichtlichkeit nur 8 der 10 Photoströme eingezeichnet

3.4.4 Die Analogspeicher und der Ladungsverstärker

Um eine Interpolation der Abszissen-Koordinate des Intensitätsmaximums des Fokalkpunktes durchführen zu können, werden die Strombeträge aller Photodetektorstreifen des Cluster in Kondensatoren integriert. Nach Beendigung der Detektionsphase werden drei der gespeicherten Werte durch einen Ladungsverstärker in eine Spannung umgewandelt und durch das Single-Slope-Verfahren digitalisiert.

Die Speicherung der integralen Strombeträge erfolgt in Kondensatoren, die im Verlaufe der Detektion durch den gespiegelten Photostrom der Phototransistoren entladen werden. In Abbildung 3.40 ist der Ausgang des Stromspiegels von Abbildung 3.39 und die Elemente zum Initialisieren und Laden des Kondensators, sowie der Ladungsverstärker dargestellt. Kern der

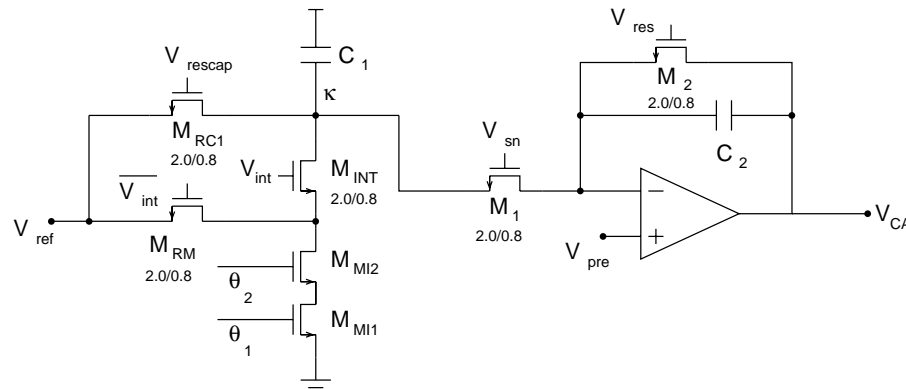


Abbildung 3.40: Analogspeicher und Ladungsverstärker des *Prototyp 2*

Schaltung sind die beiden Kondensatoren C_1 und C_2 . In C_1 wird der durch den Stromspiegel duplizierte Photostrom integriert. Das Potential des Knotens κ bestimmt die Spannung des Kondensators während der Initialisierungsphase. Durch V_{ref} unter Aktivieren von V_{rescap} kann das Potential über $M_{\text{RC}1}$ in den mittleren Spannungsbereich gelegt werden. Zu Beginn eines Integrationszyklus wird durch Aktivierung von V_{int} und Deaktivierung von V_{rescap} der Knoten κ von der Referenzspannung gelöst und der Integrationsstrom entlädt die Ladung des Kondensators C_1 , so dass das Potential des Knotens κ sinkt. Der Transistor M_{RM} hält das Drain-Potential des Stromspiegeltransistors $M_{\text{MI}2}$ während der Initialisierungsphase auf dem Potential der Spannung V_{ref} , um ein parasitäres Entladen des Knotens κ durch Angleichen des Drain-Potentials von $M_{\text{MI}2}$ zu Beginn des Integrationszyklus zu vermeiden.

Nach Beendigung des Integrationszyklus wird V_{int} deaktiviert. Anschließend wird V_{sn} durch das Schaltnetzwerk von Abbildung 3.38 aktiviert und die Ladung des Kondensators C_1 fließt zum Ausgleichen des Potentials von κ und der Referenzspannung V_{pre} unter Deaktivierung von V_{res} in den Kondensator C_2 . Die Spannung über dem Kondensator C_2 bestimmt die Ausgangsspannung V_{CA} des Ladungsverstärkers. Sie beträgt

$$V_{CA} = V_{pre} - \left(V_{ref} - \frac{C_2}{C_1} \int_0^t I_{D(MMI2)} dt - V_{pre} \right) \quad (3.53)$$

Die Referenzspannungen V_{ref} und V_{pre} werden derart gewählt, dass ein maximaler Spannungsbereich unter Berücksichtigung der Möglichkeiten des Operationsverstärkers des Ladungsverstärkers

ausgeschöpft wird. Unter Wahl von $V_{pre} = 2.5\text{ V}$ und $V_{ref} = 4\text{ V}$ beträgt die Ausgangsspannung des Operationsverstärkers je nach gespeicherter Ladungsmenge des C_1 zwischen 1 V und ca. 4 V.

In der praktischen Realisierung der Schaltung entsteht durch Mismatching eine Unsicherheit der Ausgangsspannung, die unter Umständen einige mV betragen kann. Diese Offsetspannung addiert sich aus den Beträgen des Operationsverstärkers, des Clock-Feedthrough der Transistoren M_1 , M_{RC1} und M_{INT} und aus einer Unterschiedlichkeit der Kondensatoren C_1 und C_2 . Unter Annahme einer Offset-Spannung des Operationsverstärkers von 10 mV, eines Clock-Feedthrough-Anteils²⁶ von ca. 2 mV und einem Mismatching der Kondensatoren von ca. 5 % ergibt sich eine gesamte Offset-Spannung von $\Delta V_{CA} \approx 20\text{ mV}$. Es resultiert

$$V_{CA} = 2 V_{pre} - V_{ref} + \frac{C_2}{C_1} \int_0^t I_{D(MM12)} dt \pm \frac{1}{2} \Delta V_{CA} \quad (3.54)$$

3.4.5 Das Digitalisierungsverfahren

Zur Umwandlung des analogen Spannungswertes der Ladungsverstärker wurde ein Single-Slope-Verfahren gewählt. In dem Verfahren wird die zu digitalisierende Spannung mit einer stetig fallenden Referenzspannung verglichen und die Zeitdauer bis zum Erreichen einer Gleichheit durch einen digitalen Zähler quantifiziert. Die Geschwindigkeit, mit der der Zähler arbeiten kann, bestimmt die nominale Auflösung des Verfahrens. Die Genauigkeit des Vergleichens, sowie die Linearität der Referenzspannungs-Rampe bestimmen die verwertbare Auflösung des Verfahrens (ENOB²⁷). Im Folgenden werden der Komparator und der Generator der Referenzspannungsrampe des *Prototyp 2* vorgestellt.

Der Komparator

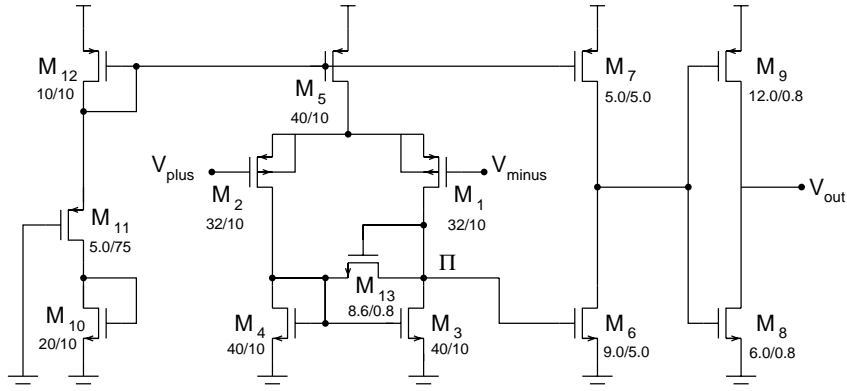
Die Zählrate des digitalen Zählers, dessen Zählerstand bei Gleichheit von Ladungsverstärker-Spannung und Referenzspannung in die entsprechenden Schieberegisterzellen übertragen wird, beträgt 2.5 MHz. Die Zeitdauer einer Zählstufe ergibt sich daher zu $t_{cnt} = 400\text{ ns}$. Um die Auflösung der 10-Bit-Architektur zu nutzen, muss die Referenzspannung innerhalb des Zeitraumes von 400 μs den vollen Spannungsbereich der Ausgangsspannung des Ladungsverstärkers von 3 V überdecken. Die geforderte Komparatorgeschwindigkeit beträgt daher 400 ns, in denen bei Gleichheit der Vergleichsspannungen das Ausgangssignal aktiviert werden muss.

Die Architektur des verwendeten Komparators ist in Abbildung 3.41 dargestellt. Es handelt sich um einen unkompenzierten Miller-Operationsverstärker mit nachgeschaltetem Inverter. Der MOS-FET M_{13} verhindert eine tiefe Sättigung des M_3 , indem er das Potential des Knotens II bei Aktivität des M_1 über den Betrag der Gate-Source-Spannung mit dem Drain-Potential des M_4 verbindet. Dadurch ist ein schnelleres Umladen der Kapazität des Knotens II möglich.

Die MOS-FET M_{10} , M_{11} und M_{12} realisieren eine Bias-Stufe, die einen Ruhestrom von ca. 10 μA generiert. Die Dimensionierung der übrigen MOS-FET erfolgte unter Abwägen der Verstärkungen und der Größe der Slew-Rate der einzelnen Stufen. Der Komparator benötigt nominal eine Verstärkung von ca. 65 dB, um ein Eingangssignal in der Größe von $V_{plus} - V_{minus} \approx \frac{3}{1024} V$ in einen Ausgangsspannungshub von ca. 5 V umzusetzen. Dabei ist gleichzeitig die Geschwindigkeit zu beachten, die sich durch das Umladen der verschiedenen Kapazitäten innerhalb des Komparators ergibt.

²⁶ Aufgrund der großen Lastkapazität von ca. 1 pF von C_1 und C_2 gegenüber den Gate-Drain Überlappkapazitäten der Pass-Transistoren von ca. 2 mV ist diese Spannung sehr gering.

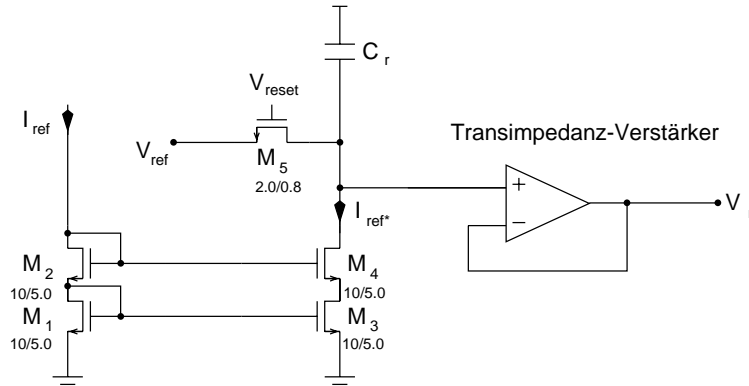
²⁷ Effective Number Of Bits

Abbildung 3.41: Der Komparator des Single-Slope-Verfahrens des *Prototyp 2*

Der Miller-Operationsverstärkers wurde derart dimensioniert, dass die Verstärkung der Ein- und Ausgangsstufe moderat und die Slew-Rate hoch ist. Der angeschlossene Inverter bewirkt einen weiteren Verstärkungsfaktor und behält die Slew-Rate des Operationsverstärkers weitestgehend bei.

Der Rampengenerator

Der Rampengenerator des *Prototyp 2* erzeugt die für das Single-Slope Verfahren benötigte Rampe. Dazu wird über eine externe programmierbare Stromquelle, deren Strom über einen Eingangspad auf den Chip geführt wird, ein Kondensator auf dem Chip entladen, so dass sich eine stetig fallende Spannung ergibt. Um die gesamte Last-Kapazität, die sich aus den insgesamt 54 Komparator-Eingängen zusammensetzt, treiben zu können, wurde ein Transimpedanz-Verstärker entwickelt. Abbildung 3.42 stellt die Konfiguration des Rampengenerators dar.

Abbildung 3.42: Der Rampengenerator des *Prototyp 2*

Durch den Strom I_{ref*} wird der Kondensator C_r , der vor Beginn einer Digitalisierung auf die Spannung $V_{dd} - V_{ref}$ aufgeladen wurde, entladen. Das Potential des nichtinvertierenden Einganges des Operationsverstärkers beträgt

$$V_{plus}(t) = V_{ref} - \frac{1}{C_r} \int_0^t I_{ref*} dt \quad (3.55)$$

Der kaskadierte Stromspiegel aus M_1 , M_2 , M_3 und M_4 spiegelt den über einen Pad des *Prototyp 2* eingeprägten Konstant-Strom I_{ref} in den Strom I_{ref}^* . Der Stromspiegel besitzt aufgrund der kaskadierten Form einen sehr hohen Ausgangswiderstand, der einen von der Ausgangsspannung nahezu unabhängigen Ausgangstrom verursacht. Die Unabhängigkeit des Stromes von der Spannung ist zu der Erzeugung einer linearen Spannungsrampe durch Entladen des Kondensators C_r erforderlich. Die Konfiguration ermöglicht eine differentielle Nichtlinearität der Spannungsrampe von weniger als 1 mV.

Der Operationsverstärker entkoppelt als Transimpedanzverstärker die Impedanz des Rampenspannungs-Netzwerkes des *Prototyp 2* von der Impedanz des positiven Eingangs-Knotens des Operationsverstärkers. In Abbildung 3.43 ist der zweistufige Aufbau des Verstärkers als Miller-Operationsverstärker mit einer Emitter-Folger-Stufe eines Bipolartransistors zur Reduktion des Ausgangswiderstandes gezeigt.

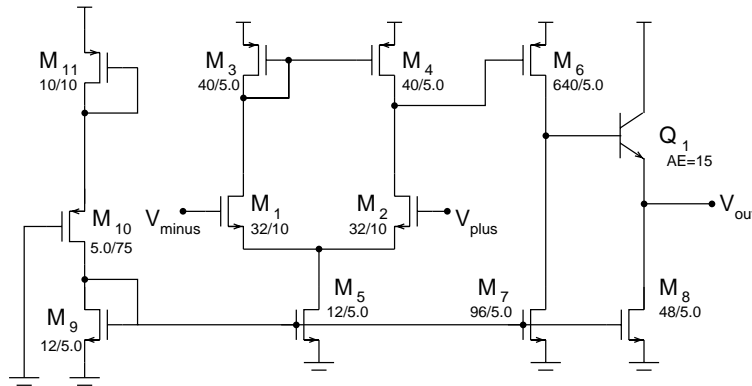


Abbildung 3.43: Der Transimpedanzverstärker des Rampengenerators

3.5 Erkenntnisse aus der Entwicklung von Prototyp 1 und 2

Die im Rahmen des Einsatzes des *Prototyp 1* als Augenbewegungs-Verfolger durchgeführten Messungen der Eigenschaften des ASIC zeigten, dass die erforderliche Lichtleistung eines Fokalkpunktes des Linsenarray mehr als 50 nW betragen musste, um mit einer Frequenz von 1 kHz eine Bestimmung der Fokalkpunkt-Position mit einer Pixel-Genauigkeit zu ermöglichen. Die in *Prototyp 1* realisierte Schaltung der WTA-Module verwendet keine Initialisierung der Anfangspotentiale, so dass die Geschwindigkeit der Detektion der Photostrom-Unterschiede von benachbarten Photodetektor-Streifen durch die erforderliche Umladung der Potentiale der Knoten Ψ der entsprechenden WTA-Zellen reduziert wurde.

Das Konzept der Initialisierung der Anfangspotentiale der Knoten Ψ wurde daher in dem *Prototyp 2* angewendet, um eine Beschleunigung des Detektionsvorganges zu ermöglichen. Gleichzeitig wurde der Algorithmus zur Interpolation der Fokalkpunkt-Positionen durch Vergleich der analogen Signalwerte von drei Pixeln eingebracht. Um weiterhin eine Erhöhung der durch die geringe Lichtleistung der Fokalkpunkte erzeugten schwachen Photoströme zu bewirken, wurde die Schaltung des ASIC in der BiCMOS-Technologie der Firma AMS prozessiert, um die vertikalen npn-Phototransistoren als Photodetektoren einzusetzen. Die erforderliche Sensitivität zur Detektion der Fokalkpunkte verbesserte sich auf 2 nW bei einer Wiederholrate von 1 kHz, um die Fokalkpunkt-Positionen durch die WTA-Module mit Pixel-Genauigkeit zu bestimmen.

Die Interpolation der Positionen durch *Prototyp 2* erbrachte jedoch keine weitere sinnvolle Verbesserung der Positionsbestimmung, da das Digitalisierungsverfahren keine verwertbaren Ergebnisse lieferte. Dies lag zum einen an der geringen realisierten Genauigkeit des Digitalisierungs-Verfahrens, das durch das Mismatching der einzelnen Kanäle der Schaltung gestört wurde, zum anderen an dem parasitären Einfluss von Photonen auf die Schaltung, da durch die Komplexität der Schaltung die beiden Metallisierungslagen des Prozesses für die Signalverarbeitung benötigt wurden.

Obwohl die Verwendung der Phototransistoren generell eine Verbesserung der optischen Sensitivität des ASIC erbrachte, wurde das Konzept des *Prototyp 2* verworfen, da die Interpolationsmöglichkeit der Positionsdetektion keine Verbesserung gegenüber der Detektionsauflösung des *Prototyp 1* ergaben. Weiterhin wurde durch die zur Interpolation erforderliche Schaltung eine große Fläche der Chip verbraucht, was die Fläche der Photodetektormatrizen relativ zu der gesamten verwendeten Fläche auf einen Anteil von weniger als 25% reduzierte.

Die Erkenntnisse aus den beiden Entwürfen wurden verwendet, um den *HSSX* zu entwickeln. Die Wahl der Phototransistoren des *Prototyp 2* erbrachte eine Verbesserung der Sensitivität des *Prototyp 2* gegenüber der des *Prototyp 1*. Die Verwendung dieser Photodetektoren erfordert jedoch die Benutzung des BiCMOS-Prozesses der Firma *AMS*, der zum einen um den Faktor von 1.5 höhere Kosten der Produktion verursacht, zum anderen beträgt die minimale Strukturgröße der Elemente $0.8\ \mu\text{m}$ und es werden nur 2 Metallisierungsebenen prozessiert. Gleichzeitig bestand die Möglichkeit, den $0.6\ \mu\text{m}$ -CMOS-Prozess zu verwenden, der aufgrund der reduzierten Strukturgrößen eine komplexere Signalverarbeitung auf gleicher Fläche ermöglicht und zusätzlich eine dritte Metallisierungsebene zum Schaltungsentwurf anbietet.

Letztendlich gab die Möglichkeit der dritten Metallisierungsebene den Ausschlag zur Verwendung des $0.6\ \mu\text{m}$ -CMOS-Prozesses, da insbesondere das parasitäre Einfallen von Photonen in den Bereichen der WTA-Module durch eine komplette Bedeckung mit einer Metall-Lage verhindert werden kann und gleichzeitig eine komplexere Signalverarbeitung möglich ist. Obwohl die spektrale Empfindlichkeit der npn-Bipolartransistoren gegenüber der der p+/nwell-Dioden in dem relevanten Bereich von 680 bis $800\ \mu\text{m}$ um einen Faktor von 10 bis 15 höher ist (siehe Abbildungen 2.7 und 2.4), wurde der Vorzug aufgrund der oben genannten Vorteile der Prozessierung mit dem $0.6\ \mu\text{m}$ -CMOS-Prozess gegeben.

In dem folgenden Kapitel werden die zur Verbindung der drei Wellenfrontsensor-ASICs mit einem Datenverarbeitungs-System entwickelten externen Systeme vorgestellt.

Kapitel 4

Das externe System

In diesem Kapitel wird der Aufbau des externen Systems zur Ansteuerung der im Projekt entwickelten ASICs erläutert. Die Daten, die die Chips während einer Detektionsphase generieren und in einer anschließenden Schiebephase an ihre Ausgangs-Pads schieben, werden durch einen FPGA weiterverarbeitet und über eine Standardschnittstelle an ein Datenverarbeitungs-System übertragen. Ziel der Entwicklung des externen Systems war es, eine flexible, kompakte und kostengünstige Konfiguration zu schaffen. Im Folgenden wird der Entwurf des FPGA und die verschiedenen Module zum Anbinden der drei Prototypen des Projektes an das externe System vorgestellt.

4.1 Systemarchitektur

Abbildung 4.1 zeigt den schematischen Aufbau des externen Systems. Zentrale Einheit ist ein FPGA der Firma *Xilinx Inc.* des Typs XC4006E. Der FPGA verbindet die Prototypen des Wellenfrontsensor-ASIC mit einem Datenverarbeitungssystem, indem er den Datentransport zwischen dem ASIC und dem RAM und zwischen dem RAM und dem Datenverarbeitungs-System steuert. Die Daten werden dabei zu einer Übertragung umstrukturiert und über das *ECP*¹-Protokoll an die parallele Schnittstelle eines Pentium-PC geleitet, der die Daten weiterverarbeitet. Im Rahmen des Projektes wurden zwei etwas unterschiedliche Konfigurationen des externen Systems entworfen. *Prototyp 1* und *Prototyp 2* übertragen die generierten Daten direkt in den FPGA, der sie in einem internen RAM zwischenspeichert, umstrukturiert und an den Parallel-Port des PC ausgibt. Das *System 1* besteht aus den Komponenten ASIC, FPGA, DAC² und PC. Der Prototyp *HSSX* benutzt zur Zwischenspeicherung der Daten einen externen RAM, da die Anzahl der Daten von 512 Byte für das interne RAM des FPGA des Typs XC4006e zu umfangreich ist. Die Daten des *HSSX* werden in einem Dual-Port-RAM der Firma *IDT Inc.* des Typs IDT7025 zwischengespeichert und unter Steuerung des FPGA an den Parallel-Port des PC ausgegeben. Das *System 2* besteht aus den Komponenten ASIC, FPGA, DAC, RAM und PC.

Der FPGA fungiert als Controller des Systems. Ihm unterliegt die Steuerung der ASICs in Bezug auf die Erzeugung der verschiedenen digitalen Steuersignale zur Generation der Positionsdaten der Fokalfunkte, die Steuerung der DA-Wandler des Systems zur Definition der analogen

¹Extended Capabilities Port

²Digital Analog Converter

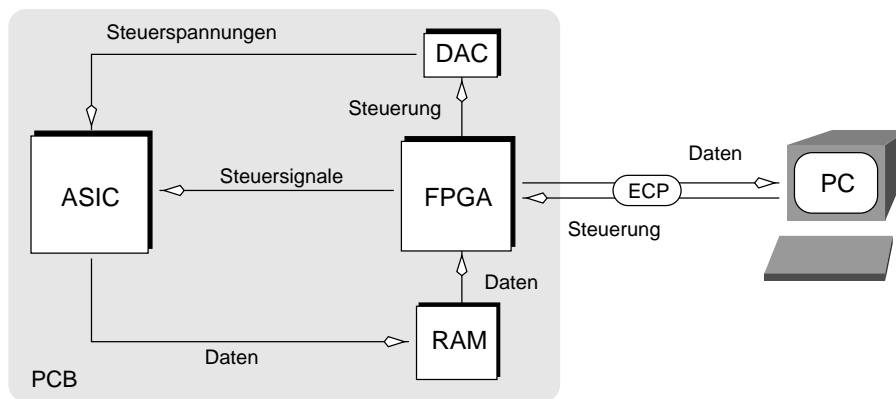


Abbildung 4.1: Die Architektur des externen Systems

Steuerspannungen der ASICs und die Kontrolle der Schnittstelle zum PC. In diesem Kapitel wird hauptsächlich die im FPGA realisierte Schaltung zur Steuerung des *HSSX* mit *System 2* erläutert. Die logischen Strukturen des FPGA zur Steuerung des *System 1* mit den beiden ersten Prototypen sind bis auf die Steuerung des RAMs funktionell identisch.

Das System 1

In Abbildung 4.2 sind die PCBs des *System 1* dargestellt. Es handelt sich um eine im oberen Teil der Abbildung dargestellte Basisplatine, an die über Steckverbinder wahlweise das PCB des *Prototyp 1* oder des *Prototyp 2* angeschlossen werden kann. Die beiden Prototypen wurden direkt auf PCBs kontaktiert, da eine Montage der Chips an optomechanischen Systemen vorgesehen war und die typischen Chipcarrier aus Gründen der Justage und Anpassung nicht eingesetzt werden konnten.

Auf der Basisplatine ist ein vierfacher 8-Bit-DAC angeordnet, um die analogen Bias-Spannungen von *Prototyp 1* und *Prototyp 2* zu erzeugen. Der 12-Bit-DAC wird zur Erzeugung eines Referenzstromes benötigt, der die Rampenspannung des Single-Slope-Digitalisierungsverfahren von *Prototyp 2* generiert. Zur Definition einer präzisen Referenzspannung des 12-Bit-DACs ist ein diskreter Bandgap-Generator vorhanden, der eine Referenzspannung von 2.5 V liefert. Der EEPROM wird zum Laden der Konfigurations-Daten des FPGA verwendet.

Das System 2

In Abbildung 4.3 ist das PCB des System 2 dargestellt. Die drei PLCC³-Gehäuse des FPGA, des DP-RAM und des *HSSX* treten augenscheinlich hervor. Im linken unteren Teil sind die Digital-Analog-Wandler zur Erzeugung der analogen Bias- und Steuerspannungen des *HSSX* und die Präzisions-Spannungswandler zur Erzeugung einer stabilen 5 V und 10 V-Versorgungsspannung des Systems zu sehen. Letztere wird für die Erzeugung einer Spannung V_{bulk} benötigt, die in einem Bereich von 5 bis 7 V eingestellt werden kann. Dazu wird die Ausgangsspannung eines DAC-Kanals mit einem Spannungsbereich von 0 bis 2.5 V durch den Präzisions-Operationsverstärker (*Opamp*) mit einem festen Verhältnis verstärkt. Bei den DACs handelt es sich um vierfache 10-Bit-Wandler, die durch eine diskrete Referenzspannungsquelle von 2.5 V versorgt werden.

Die Anordnung der Bauelemente auf der Platine erfolgte unter Trennung der analogen und digitalen Funktionseinheiten, um das Übersprechen der digitalen Spannungen auf die sensiblen

³Plastic Leaded Chip Carrier

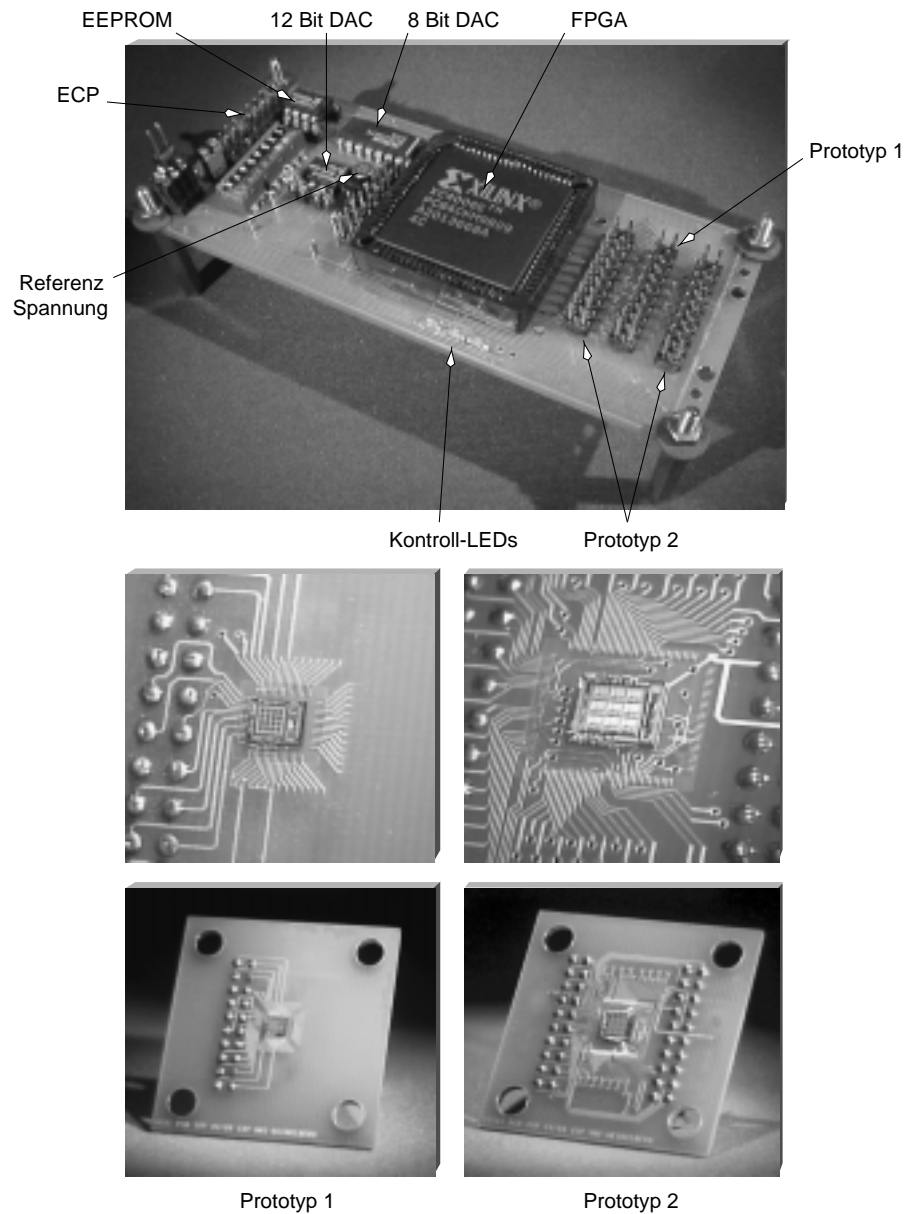


Abbildung 4.2: Das *System 1* zur Verbindung der beiden ersten Prototypen mit einem PC. Die beiden Prototypen wurden zur Montage an optomechanischen Komponenten direkt auf PCBs kontaktiert.

analogen Steuerspannungen des *HSSX* zu reduzieren. Um während den in Kapitel 6 beschriebenen Messungen Einflüsse des Übersprechen vollständig zu unterbinden, erfolgte die Detektion der Fokuspunkte-Positionen und das Auslesen der Daten an den ECP in einer zeitlichen Abfolge.

4.2 Die Schaltung des FPGA

Der FPGA realisiert die Verbindung zwischen dem Datenverarbeitungs-System (im Folgenden als PC bezeichnet) und dem Wellenfrontsensor-ASIC. Er erzeugt zum einen die Steuersignale des ASIC, um die Abfolge der verschiedenen Phasen wie Detektion, Laden der Schieberegister,

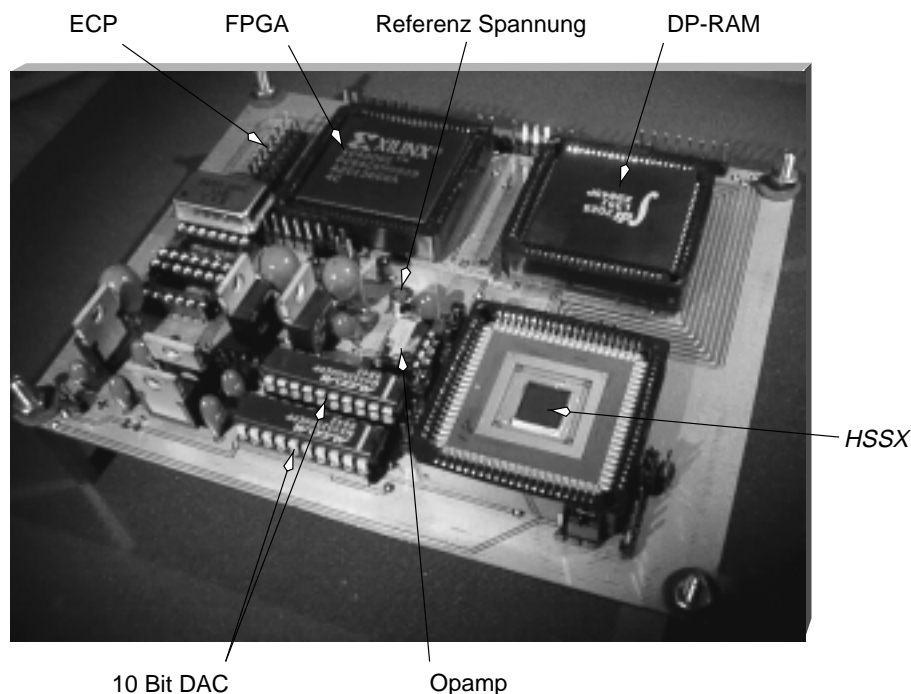


Abbildung 4.3: Das *System 2* mit dem *HSSX*

Schieben und gleichzeitiges Komprimieren der Daten zu regeln. Zum anderen anderen kontrolliert er die Übersendung der vom ASIC erzeugten und im RAM zwischengespeicherten Daten an den PC. Um diese Funktionen logisch zu strukturieren, unterteilt sich die Schaltung des FPGA in mehrere Bereiche. Der eine Bereich enthält die Steuerlogik der ECP-Schnittstelle zur Kontrolle einer bidirektionalen und asynchronen Datenübertragung. Dieses Modul wird aufgrund seiner Funktion als *Interface* bezeichnet. Der zweite Bereich umfasst eine als *Cycler* bezeichnete Steuerlogik, die zyklisch eine Abfolge von Steuersignalen erzeugt, um den Datentransfer zwischen ASIC und RAM zu regeln. Die beiden Schaltungsteile werden durch ein drittes Modul in Form eines globalen Arbitrierungsnetzwerkes innerhalb des FPGA miteinander verbunden und überwacht. Dieses zentrale Modul wird als *Controller* bezeichnet. Der Datentransfer zwischen dem RAM, dem ASIC und dem Interface wird durch die Einheit *Ramcontrol* übernommen. Ein weiteres Modul *Loaddac* ist für die Ansteuerung der verschiedenen DA-Wandler von *System 1* und *System 2* zuständig.

Die Anwender-Software des PC übergibt über die parallele Schnittstelle Steuerkommandos an den Controller, der aufgrund der Kommandos verschiedene interne und externe Signale erzeugt, mit denen die einzelnen Funktionsgruppen des FPGA gesteuert werden. Der FPGA verfügt über zwei Betriebsarten, die im Folgenden beschrieben werden.

In der ersten Betriebsart, bezeichnet als *Single Frame*, erfolgt das Auslesen einer einzelnen Bildinformation. Die kompletten Daten einer Positions-Detektion des ASIC werden in das RAM eingelesen und danach an die parallele Schnittstelle des ASIC übertragen. Anschließend befindet sich der FPGA in einem Warte-Zustand.

Die zweite Betriebsart, bezeichnet als *Multiple Frame*, ermöglicht das mehrmalige Auslesen der Positionsdetektions-Daten aus dem RAM. Der Cycler des FPGA erzeugt wiederholt die zyklische Abfolge der Steuersignale, um den ASIC Bilddaten erzeugen zu lassen und die Daten zyklisch in das RAM zu schreiben. Es ist möglich, über den PC ein einzelnes Bild aus dem RAM auszulesen,

oder synchron in Echtzeit die Daten der repetitiven Bilder aus dem RAM auszulesen. Im ersten, asynchronen Fall übernimmt der PC als sogenannter *Master* die Kontrolle über das Verhalten des FPGA, der einen *Slave*-Status besitzt. Im zweiten, synchronen Fall wird der FPGA als Master definiert und der PC erwartet in seiner Funktion als Slave eine Information, die ihm das vollständige Schreiben eines Bildes in das RAM anzeigt.

Der PC ist in diesem Betriebsmodus als Slave über den FPGA gesteuert, um ein zu den Detektionsphasen synchrones Auslesen der Daten aus dem RAM des Systems zu erlauben. Diese Funktion erfordert in dem verwendeten Betriebssystem *Windows 95* eine Maskierung der betriebssystemeigenen Interrupts, um einen Kontextwechsel⁴ zu verhindern. Um bei Ausbleiben einer Antwort des FPGA die Betriebsfähigkeit des PC zu sichern, verfügt die Anwender-Software des PC über einen Kontrollmechanismus, der bei einer Zeitüberschreitung die Interruptsteuerung demaskiert.

Abbildung 4.4 stellt die einzelnen Module des FPGA mit ihren wichtigsten Steuersignalen dar. Die Richtung der Signalpfade ist durch einen Pfeil gekennzeichnet. Die punktierten Signalpfade sind interne, die durchgezogenen Signalpfade externe Signale. Die hervorgehobenen Signalpfade stellen die internen Handshake-Signale der einzelnen Module dar. Im Folgenden wird zuerst die Funktion des Interface unter Betrachtung des ECP-Protokolls erläutert. Nachfolgend werden der Controller sowie alle anderen Module des FPGA beschrieben.

Die logischen Schaltungen des FPGA wurden durch die Hardware-Beschreibungs-Sprache *Verilog* beschrieben und durch das Synthese-Programm *Synopsis* der Firma *Synopsis Inc.* in die Abstraktionsebene der Gatter-Bausteine übersetzt. Die anschließende Programmierung des FPGA erfolgte mit Hilfe des durch die Software *Design Manager* von *Xilinx* generierten Bitstreams. Alle Module wurden in Form einer vollsynchronen Zustandsmaschine entworfen und besitzen ein asynchrones Reset zur Definition eines Anfangszustandes. Die Systeme verwenden einen nichtüberlappenden Zweiphasentakt mit den beiden Taktsignalen Φ_1 und Φ_2 . Im Allgemeinen werden die Zustände der Zustandsmaschinen bei einer positiven Taktflanke des Signals Φ_2 gewechselt.

Interface

Das Interface realisiert das zum Datentransfer über die parallele Schnittstelle erforderliche ECP-Protokoll. Der ECP-Baustein des PC verfügt über eine 16 Byte tiefe FIFO⁵-Architektur, deren Datenaustausch mit einem Peripherie-Gerät durch ein Hardware-Protokoll gesteuert wird. Die Anwender-Software schreibt Daten in die FIFO und überprüft ein Kontrollregister des Port-Bausteins auf das vollständige Entleeren der FIFO durch das Peripherie-Gerät. Durch den mit Hardware-Handshake gesteuerten Datenaustausch werden nach [Microsoft 93] Transfer-Raten von bis zu 2 MByte/s ermöglicht.

Im Betrieb überprüft das Interface die Steuersignale des Parallel-Ports des PC und übernimmt nach Aktivierung des Steuersignals *HostClk* durch ein Handshake-Verfahren mit dem Signal *PeriphAck* genau drei Bytes. Diese drei Bytes stellen ein Kommando-Byte sowie zwei Daten-Bytes als Argumente des Kommando-Byte dar. Nach erfolgreichem Abschließen der Übernahme der drei Bytes wird ein internes Handshake-Signal aktiviert, das dem Controller eine gültige Befehlssequenz signalisiert. Das Interface wartet auf die Antwort des Controllers, der den Befehl abarbeitet und hält solange das Antwort-Signal für den Parallel-Port in einem inaktivem Zustand. Dadurch wird dem PC das momentane Abarbeiten eines Befehls signalisiert. Nach erfolgtem Signal des Controllers aktiviert das Interface das Antwort-Signal und geht in den Warte-Zustand

⁴Unterbrechung des Programmflusses durch einen Interrput zur Abarbeitung der entsprechenden Anfrage

⁵First In First Out

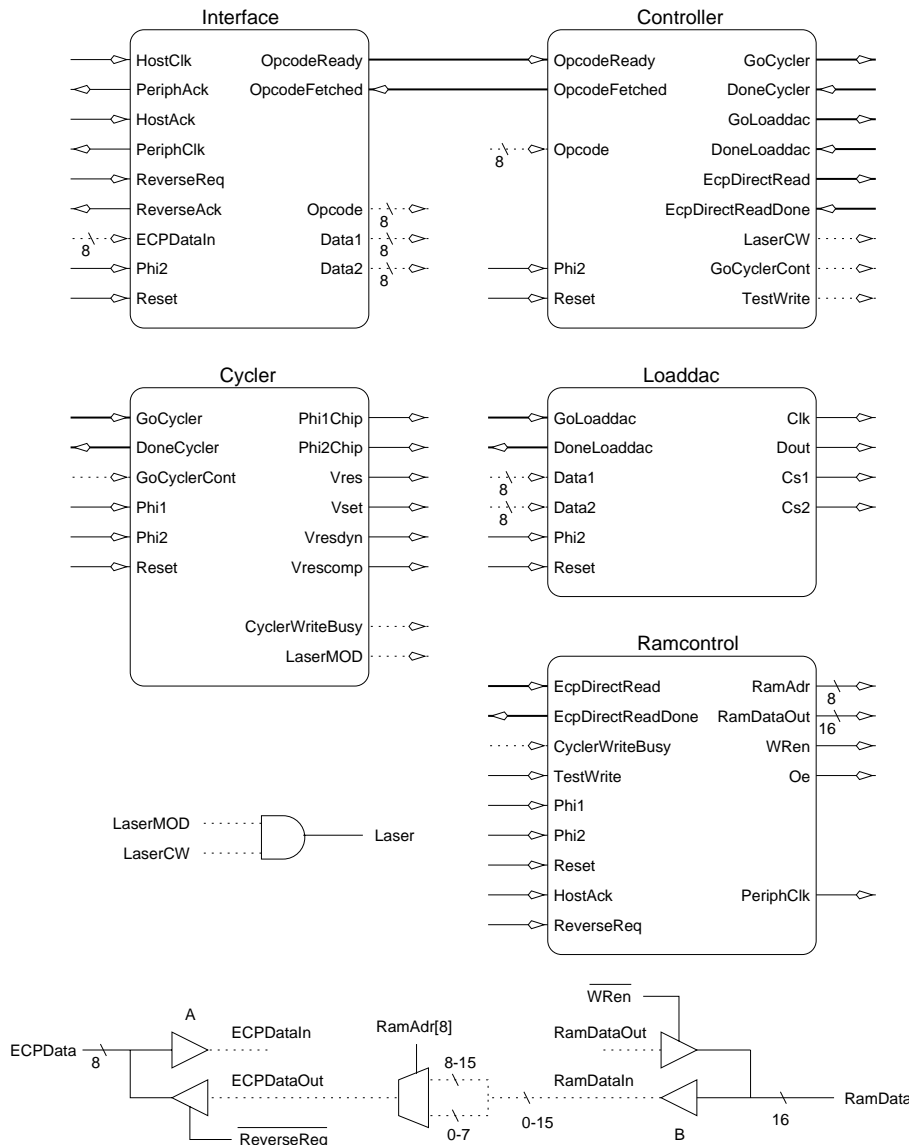


Abbildung 4.4: Die Module des FPGA

über. Abbildung 4.5 stellt den Zyklus zur Übergabe der drei Bytes und zum Aktivieren des Controllers dar.

Zu Beginn einer Datenübertragung des Parallel-Ports zieht der PC-eigene Schnittstellen-Chip das Signal *HostClk* nach logisch 0. Das Interface übernimmt an der nächstfolgenden positiven Taktflanke von Φ_2 das Datenbyte, das der Schnittstellen-Chip des PC auf den Datenleitungen *ECPData*[7:0] zur Verfügung stellt, in das FPGA-eigene Register *Opcode*[7:0] und aktiviert das Signal *PeriphAck* nach logisch 1. Der PC quittiert mit einer logischen 1 von *HostClk* und das Interface legt nach erfolgtem Quittieren synchron zu Φ_2 das Signal *PeriphAck* auf logisch 0. Danach signalisiert der PC das Übertragen eines weiteren Bytes durch das Signal *HostClk*. Dieser Vorgang wiederholt sich noch zweimal, wobei die Bytes der Datenleitung *ECPData*[7:0] nacheinander in die FPGA-eigenen Register *Data1*[7:0] und *Data2*[7:0] geschrieben werden.

Nach dem letzten Byte hält das Interface das Signal *PeriphAck* bei logisch 0 und aktiviert das Signal *OpcodeReady*, um dem Controller das Vorliegen einer gültigen Befehlssequenz anzuzeigen.

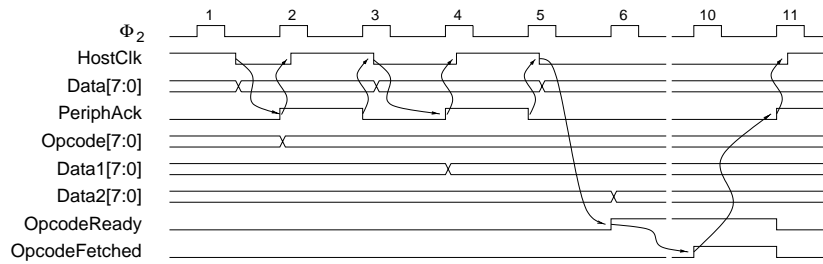


Abbildung 4.5: Das Handshake des Interface im ECP-Protokoll

Der Controller antwortet nach einiger Zeit, in der der entsprechende Befehl abgearbeitet wurde, mit dem Signal *OpcodeFetched* und das Interface zieht *PeriphAck* nach logisch 1. Dadurch ist der Vorgang abgeschlossen und das Interface geht in den Warte-Zustand über.

Controller

Der Controller empfängt das Signal *OpcodeReady* des Interface-Moduls und überprüft den Inhalt des Registers *Opcode[7:0]*. Der Inhalt von *Opcode* wird durch die Anwender-Software des PC definiert und repräsentiert einen Befehl, den der FPGA ausführen soll. Entsprechend dem Inhalt von *Opcode[7:0]* werden entweder die Handshake-Signale *GoCycler*, *GoLoaddac* oder *EcpDirectRead* zur Aktivierung des entsprechenden Moduls angesteuert oder interne Status-Register des FPGA gesetzt. Im ersten Fall der Aktivierung eines Moduls wartet der Controller auf das Antwortsignal *DoneCycler*, *DoneLoaddac* oder *EcpDirectReadDone* des entsprechenden Moduls. Nach Empfang des Antwortsignals aktiviert der Controller die Leitung *OpcodeFetched* und geht in einen Wartezustand über. Im zweiten Fall setzt der Controller die internen Register entsprechend der Register *Data1[7:0]* und *Data2[7:0]* und erzeugt sofort das Signal *OpcodeFetched*.

Cycler

Das Modul Cycler erzeugt die zeitliche Abfolge von Signalen, die zur Ansteuerung der einzelnen Module des angeschlossenen ASIC erforderlich sind, um die Informationen einer Fokalkpunkt-Positionsdetektion zu erzeugen. Im Falle des *Prototyp 1* wird das Set-Signal gesetzt, um die Daten der WTA-Zellen in die Schieberegister zu übertragen und anschließend wird eine Abfolge von 32 Taktzyklen der Takte Φ_1 und Φ_2 zum Schieben der Daten generiert. Zur Ansteuerung des *HSSX* werden die Signale V_{set} , V_{res} , V_{resdyn} , V_{rescomp} , Φ_1 und Φ_2 gemäß der Abbildung 3.31 erzeugt. Das Modul besitzt einen synchronen Zähler, dessen Zählerstände mit dem Wert von beschreibbaren Registern verglichen werden. Bei Gleichheit wird das entsprechende Signal aktiviert oder deaktiviert. Durch die flexible Implementierung der Vergleichswerte ist es möglich, die Zeitkonstanten der Zeit-Diagramme der ASICs im Betrieb durch die Anwender-Software des PC einzustellen und bei Bedarf zu ändern. Das Signal LaserMOD erzeugt ein pulswidenmoduliertes Signal zur Steuerung eines Lasers, der bei Bedarf synchron zu der jeweiligen Detektionsphase der Fokalkpunkt-Positionen aktiviert wird.

Das Modul besitzt zwei Betriebsarten. Im *Single Frame*-Modus erhält es das Start-Signal *GoCycler* des Controller-Moduls und durchläuft genau einen Zyklus zur Generation der Steuer-Signale. Danach erzeugt es das Signal *DoneCycler* und geht in einen Warte-Zustand über. Im *Multiple Frame*-Modus, der durch das Signal *GoCyclerCont* des Controllers angezeigt wird, läuft der Zähler des Modules ununterbrochen, bis das Signal *GoCyclerCont* durch den Controller deaktiviert wird.

Loaddac

Das Modul Loaddac erzeugt einen Bitstring, um die analoge Ausgangsspannung der DA-Wandler des PCBs zu setzen. Dazu werden die Daten der Register *Data1* und *Data2* entsprechend der Definition in [Maxim] zusammengefasst und in ein 16-Bit-Register übertragen. Anschließend wird der Inhalt des Registers zählergesteuert sequentiell an den Port *Dout* übergeben. Gleichzeitig wird das Taktsignal *Clk* und das Chip-Select-Signal *Cs1* oder *Cs2* erzeugt. Das Signal *GoLoaddac* des Moduls Controller startet diesen Vorgang. Nach Beendigung aktiviert das Modul Loaddac das Signal *DoneLoaddac*.

Das DP-RAM

Abbildung 4.6 stellt die Anbindung des DP-RAM an den ASIC und den FPGA dar. Die Informationen *OutX[7:0]* und *OutY[7:0]* des *HSSX* können aufgrund einer 16-Bit Datenbreite des DP-RAM parallel in dessen Speicher geschrieben werden. Das Auslesen der Daten durch Ramcontrol erfolgt für die zwei Dimensionen der Positions-Daten des *HSSX* sequentiell, da die Datenbreite der Parallel-Port-Verbindung auf 8 Bit begrenzt ist. Es werden nacheinander die Daten der X-Richtungen und anschließend die der Y-Richtungen ausgelesen. In der Abbildung sind die

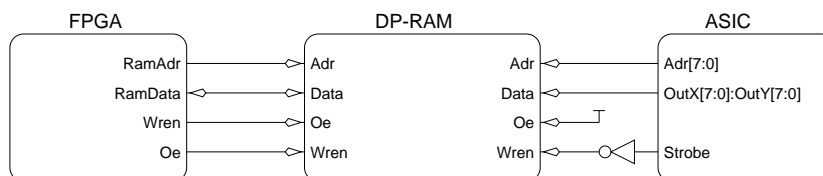


Abbildung 4.6: Das Dual-Port RAM als Zwischenspeicher

relevanten Verbindungen dargestellt. Das DP-RAM besitzt einen Speicher mit einer Tiefe von 16 KByte, von dem 512 Byte verwendet werden. Weiterhin verfügt es über eine semaphorische Arbitrierungslogik und Interrupt-Signale, die von dem System jedoch nicht benutzt werden, da Schreibzugriffe nur über einen der beiden Ports erfolgen und keine Konflikte bei einem gleichzeitigen Beschreiben derselben Speicherzelle auftreten. Die entsprechenden Eingangs-Signale des RAM werden im System statisch mit V_{dd} oder Masse verbunden, die entsprechenden Ausgangssignale bleiben ungenutzt.

Ramcontrol

Das Modul Ramcontrol steuert die Übertragung der Positions-Daten aus dem DP-RAM an den PC. Die Daten des *HSSX*, die während des Schiebe- und Komprimiervorganges an dessen Ausgangs-Ports anliegen, werden eigenständig durch den *HSSX* in das RAM übertragen. In *System 1* wird das Übertragen der Positions-Daten von *Prototyp 1* und *Prototyp 2* in das interne RAM des FPGA ebenfalls durch das Modul Ramcontrol gesteuert.

Nach Beendigung eines Schreibvorganges werden die Daten durch den Parallel-Port des PC ausgelesen und in der Anwender-Software weiterverarbeitet. Die Verwendung eines Dual-Port-RAM zur Zwischenspeicherung der Positions-Daten vereinfacht die Architektur der Ausgangstreiber des *HSSX*, da keine Tri-State-Funktionalität⁶ benötigt wird, wie im Falle eines Standard-RAMs.

Der *HSSX* kann die komprimierten Positionsdaten durch Steuerung der Leitungen *Adr[7:0]* und *Strobe* (siehe Abbildung 4.6) über den ersten Port des DP-RAM in dessen Speicher schreiben.

⁶Umschaltmöglichkeit zwischen einem Zustand hoher Impedanz oder logischer Signalpegel

Das Auslesen des Speichers geschieht unter Kontrolle des Moduls Ramcontrol. Wenn die Anfrage zum Auslesen der Daten des DP-RAM durch eine Befehlssequenz des PC im Controller erkannt wird, aktiviert dieser das Signal *ECPDirectRead* und startet das Modul Ramcontrol. Dieses erzeugt sofort das Signal *ECPDirectReadDone*, um dem Controller die aktive Kontrolle des FPGA während des Datentransfers zu ermöglichen. Im *Multiple Frame*-Modus überprüft Ramcontrol das Signal *CyclerWriteBusy* des Controllers, das einen aktuellen Schreibvorgang des *HSSX* zu erkennen gibt, um eine Überlappung von Lesen und Schreiben zu verhindern⁷. Solange das Signal aktiv ist, wartet das Modul Ramcontrol, um anschließend die neu geschriebenen Daten über den zweiten Port des DP-RAM auszulesen und über die internen bidirektionalen Tristate-Treiber B für den DP-RAM an die internen bidirektionalen Tristate-Treiber A für den Parallel-Port unter Kontrolle von *PeriphClk* und *HostAck* zu übergeben. Im *Single Frame*-Modus entfällt das Erwarten eines deaktivierten *CyclerWriteBusy*-Signals. Im Folgenden wird der Datenaustausch zwischen dem RAM und dem Parallel-Port dargestellt.

Zur Übertragung der Daten aus dem DP-RAM an den Parallel-Port des PC sendet der PC eine Befehlssequenz aus drei Bytes an den FPGA. Dieser erkennt die Anfrage nach einer Datenübertragung und übergibt die Kontrolle an das Modul Ramcontrol. Anschließend wird der Parallel-Port des PC durch die Software in den Empfangsmodus geschaltet. Dies geschieht anhand einer Initialisierung der Status-Register des ECP-Bausteines. Der PC aktiviert das Signal *ReverseRequest* nach logisch 0 und wartet auf die Kenntnisnahme des FPGA durch dessen Aktivierung des Signals *ReverseAck*. Danach zeigt der PC die Bereitschaft, Daten zu empfangen, durch eine logische 0 von *HostAck* an. Der nun beginnende Zyklus zum Auslesen der Daten aus dem RAM ist in Abbildung 4.7 dargestellt.

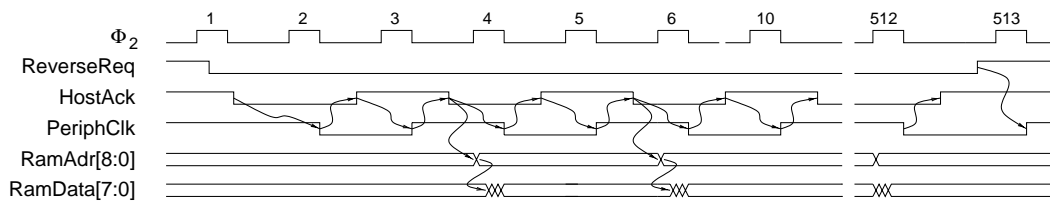


Abbildung 4.7: Das Handshake zum Übertragen der DP-RAM-Daten an den PC

Nachdem das Modul Ramcontrol die Aktivierung des Signals *HostAck* erkannt hat, zieht es das Signal *PeriphClk* nach logisch 0. Im ersten Durchgang wird der Adress-Zähler des DP-RAM nicht inkrementiert, da die 0. Speicheradresse schon vor der Initialisierung des Zyklus ausgewählt wurde und die Daten bereits anliegen. Der PC erkennt durch das Signal *PeriphClk*, dass gültige Daten anliegen, und quittiert mit *HostAck* als logisch 1. Bei der nächsten negativen Taktflanke von Φ_2 wird *PeriphClk* auf logisch 1 gehoben und der PC antwortet mit einer logischen 0 von *HostAck* um weiter Daten zu empfangen. Bei der nächstfolgenden positiven Taktflanke von Φ_2 wird der Adress-Zähler inkrementiert und ca. 25 ns später, entsprechend der Geschwindigkeit des DP-RAM, liegen gültige Daten auf dem Datenbus des DP-RAM. Das Modul Ramcontrol zieht *PeriphClk* an der nächsten negativen Taktflanke von Φ_2 nach logisch 0, um die Gültigkeit der Daten anzuzeigen.

Dieser Vorgang wiederholt sich für die 512 Byte der Positionsdaten des *HSSX*. Aufgrund der 16-Bit Datenbreite des DP-RAM und der 8-Bit Datenbreite des Parallel-Ports wird das Adressbit *RamAdr[8]* als Steuersignal eines Multiplexers verwendet, der entweder die unteren oder die

⁷Bei Benutzung eines Dual-Port-RAMs entsteht kein Konflikt bei dem Lesen einer RAM-Zelle, die zeitgleich beschrieben wird. Die Daten, die gelesen werden, sind im Allgemeinen nicht gültig. Aufgrund der zeitlichen Abfolge von Lese- und Schreib-Zugriff ist dieser Vorgang in dem System jedoch verhindert.

oberen 8 Bit des Datenbus des DP-RAM auf die Tristate-Treiber des ECP-Datenbus legt. Nach dem 512. Durchgang schaltet die Software den Parallel-Port in den Schreibmodus, wodurch das Signal *ReversReq* wieder deaktiviert wird. Das System befindet sich der nach Deaktivierung von *PeriphClk* in einem Wartezustand.

Das Modul Ramcontrol ermöglicht weiterhin einen Testmodus, in dem die einzelnen Speicherzellen des DP-RAM beschrieben werden können. In dem Testmodus, angezeigt durch das Signal *Testwrite*, wird ein Datenbyte des Registers *Data2* gleichzeitig in die oberen und unteren 8 Bit einer Adresse des DP-RAMs geschrieben, die in *Data1* repräsentiert ist. Dazu schaltet das Modul über das Signal *Oe* die Ausgangstreiber des DP-RAM in einen hochohmigen Zustand und aktiviert *WRen*. Dadurch wird der bidirektionale Tristate-Treiber B in den Schreibmodus versetzt und die Daten von *RamDataOut* in das DP-RAM übertragen.

4.3 Performance der Systeme

Die beiden entwickelten Systeme verwenden einen nichtüberlappenden Zweiphasentakt mit einer Frequenz von 2.5 MHz, mit dem die synchronen Zustandsmaschinen des FPGA gesteuert werden. Die Taktisgnale der Schieberegister und des Kompressors der Wellenfrontsensor-ASICs werden aus diesem Zweiphasentakt generiert. Das Auslesen der Daten der ASICs geschieht daher ebenfalls mit einer Geschwindigkeit von 2.5 MHz.

Während für die beiden Prototypen 1 und 2 der Systemtakt aufgrund der geringeren Datenmenge ausreicht, um eine im Vergleich zu der Detektionsphase kurze Zeitdauer des Schiebens zu ermöglichen, würde die Implementation eines Systemtaktes mit in Kapitel 3 definierten 32 MHz Frequenz zur Übertragung aller Informationen der Cluster-Matrix in den DP-RAM innerhalb 10 μ s einen verbesserten Aufbau der Platine von *System 2* erfordern. Bei einem Systemtakt von 2.5 MHz wird eine Zeitdauer von 136 μ s benötigt⁸, um die Positionsdaten in das RAM zu übertragen. Die innerhalb der für einen Detektionsvorgang angesetzten Zeitdauer von 1 ms verbleibende Zeit zur Integration der Photoströme reduziert sich daher auf 864 μ s und verringert so die Diskriminierungsgenauigkeit der WTA-Module.

Das *System 2* wurde zur Vermessung des *HSSX* aus Kostengründen als Teststudie unter Verwendung des Labor-eigenen Fräsbohr-Plotters hergestellt und unterliegt den entsprechenden Einschränkungen von Leiterbahn-Abständen und Leitfähigkeiten der Durchkontaktierungen. Ein industriell gefertigtes PCB ermöglicht die entsprechende Verbesserung der Performance des *System 2*, um die Übertragung der Daten des ASIC in das DP-RAM mit einer Frequenz von 32 MHz zu ermöglichen und parallel zu der nächsten Detektionsphase über den Parallel-Port auszulesen. Die Tests des *HSSX* in Kapitel 6 fanden derart statt, dass die Daten nach der Detektionsphase mit einer zeitlichen Länge von 1 ms innerhalb von 136 μ s in das DP-RAM geschrieben und anschließend innerhalb von ca. 1 ms über den Parallel-Port an den PC übertragen wurden. Prinzipiell ist es möglich, die Detektion der Fokalkpunkt-Positionen und das Übertragen der Positionsdaten der vorangehenden Detektion an den Parallel-Port zeitgleich durchzuführen. Aufgrund des möglichen Übersprechens der Funktionsgruppen auf dem PCB des *System 2* ist diese Betriebsform zur genauen Messung der Funktionalität des *HSSX* nicht gewählt worden. Die Ergebnisse der Messungen des *HSSX* unter Verwendung des *System 2* werden in Kapitel 6 vorgestellt.

Das *System 1* wurde zur Überprüfung der Funktionsfähigkeit von *Prototyp 1* und *2* verwendet. Im Kapitel 5 erfolgt eine Zusammenfassung der Teststudie eines Eye-Trackers, der unter Verwendung von *System 1* mit *Prototyp 1* aufgebaut wurde.

⁸340 Schreibzugriffe mit einer Frequenz von 2.5 MHz

Kapitel 5

Anwendung des ersten Prototyps als Eye-Tracker

Die Möglichkeit des Wellenfrontsensor-ASIC, die Wellenfront eines Lichtstrahls zu quantifizieren, eröffnet auch weitere Anwendungsgebiete als die in Kapitel 1.3 vorgestellte Messung der optischen Aberrationen des Auges. Der erste Prototyp des Wellenfrontsensor-ASIC wurde in einer Machbarkeits-Studie als Sensor eines Augenbewegungs-Verfolgers eingesetzt. Ziel des Einsatzes war die Überwachung der Augenbewegung in einem System der photorefraktiven Keratektomie.

5.1 Photorefraktive Keratektomie

Die Korrektur der Fehlsichtigkeit des menschlichen Auges kann unter verschiedenen Möglichkeiten erfolgen. Moderne Systeme der photorefraktiven Keratektomie (PRK) verwenden das Laserlicht eines gepulsten Excimer-Lasers mit einer Wellenlänge von 193 nm, um durch Applizieren eines hochenergetischen Photonenpulses die Bindungsbrücken der Moleküle der Hornhaut aufzubrechen, Material zu lösen und die topographische Struktur der Hornhaut zu verändern. Die verschiedenen Formen von Fehlsichtigkeit des menschliche Auges lassen sich durch Veränderung der Krümmung der Hornhaut korrigieren. Typische Formen von Weit- oder Kurzsichtigkeiten können durch die Abtragung von konzentrischen Kreisen verschiedener Tiefe korrigiert werden, in dem ein System von Blenden in den Ablations-Strahlengang eingebracht wird. Gegenstand der Forschung und Entwicklung sind Laser-Systeme mit *Flying-Spot*-Technik, die anstelle der durch Blenden maskierten Abtragung von großen Hornhautbereichen an einem wählbaren Ort der Hornhaut (Spot) einen Radius der Ablation von wenigen μm erzeugen. Durch laterales und axiales Scannen über der Hornhaut ist es möglich, eine beliebige Gestalt der Hornhaut zu erzeugen, um auch unsymmetrische Aberrationen wie Koma oder sphärische Aberration kompensieren zu können.

Nach [Seiler 98] ist die photorefraktive Keratektomie (PRK) zur Myopiekorrektur wissenschaftlich anerkannt für Aberrations-Kompensationen bis zu -6 dpt. Die refraktive Erfolgsrate liegt bei 90%, die Komplikationsrate unter 1%. Korrekturen von Hyperopie oder Astigmatismus sind aufgrund der an unterschiedlichen Stelle verschieden starken Gewebeschichten kritischer und unterliegen einer refraktiven Erfolgsrate von 80% bei einer Komplikationsrate bis zu 5%.

Die Behandlung des Patienten erfolgt im Allgemeinen ambulant. Die Dauer des Eingriffs beträgt ca. 10-40 s, in denen der Laserstrahl mit einer Rate von bis zu 10 Pulsen in einer Sekunde Gewebe der Hornhaut abträgt. Das behandelte Auge darf während dieser Zeitdauer nur äußerst geringe rotatorische und translatorische Bewegungen ausführen, da eine größere Fehlstellung der Hornhaut die Qualität der Ablation vermindern würde. Im Allgemeinen werden durch einen Laser-Puls Schichten von ca. 2-5 μm abgetragen, so dass eine fehlpositionierte Ablation über einer längeren Zeitdauer unerwünschte optische Aberrationen bewirkt.

Die Bewegung des Auges wird daher in den Systemen durch elektronische Bildverarbeitung überprüft. Bei einer Fehlstellung des zu behandelnden Auges über eine definierte Distanz wird ein Schutzmechanismus ausgelöst, der den Ablationsstrahl unterbricht. Die zur Zeit eingesetzten Systeme ermöglichen eine Wiederholfrequenz der Augenbewegungs-Detektion von 50 Hz, die der in Kapitel 3 erläuterten Frequenz von CCD-Systemen entspricht.

In den Laser-Systemen mit Flying-Spot-Technologie werden Scan-Frequenzen von 200 Hz verwendet, um in einer vertretbaren Zeitdauer die Aberrations-Kompensation der Hornhaut durchzuführen. Aufgrund der reduzierten Spot-Größe der Systeme ist ein mehrfaches Scannen der Hornhaut erforderlich, um die gleichen Resultate wie die Blenden-Systeme zu realisieren. Die Bewegungskontrolle des Auges muss daher in einer höheren Frequenz erfolgen, die durch die etablierten Augenbewegungs-Verfolger in Form von CCD-Kamera gestützten Bildverarbeitungssystemen nur unter erheblichem Aufwand realisierbar sind. Der Wellenfrontsensor-ASIC *Prototyp 1* wurde in einem System der PRK zur Kontrolle der Augenbewegung mit der Frequenz von 1 kHz erprobt. Im Folgenden werden die physikalischen Grundlagen der Bewegungs-Detektion erläutert und das aufgebaute System vorgestellt. Die durch *Prototyp 1* erreichbare Auflösung in der Bestimmung der Augenbewegung wird quantifiziert und Messergebnisse dargestellt. Abschließend erfolgt eine Abschätzung der Auflösung und des dynamischen Bereichs der Bewegungsverfolgung unter Verwendung verschiedener Linsenarrays.

5.2 Augenbewegungs-Detektion durch das Hartmann-Shack-Verfahren

Die Bewegungen des Auges lassen sich durch die Detektion einer unter bestimmten Voraussetzungen reflektierten Wellenfront eines einfallenden Lichtstrahles messen. Die Wellenfront wird durch das Linsenarray und den Wellenfrontsensor-ASIC erfasst und die Daten der Fokalfpunkte-Abweichungen an das externe System übergeben. Dieses ermittelt aus den Informationen die globale Verkippung der Wellenfront, die ein Maß für die rotatorische oder translatorische Bewegung des Auges ist.

Der einfallende Mess-Lichtstrahl wird auf den Mittelpunkt des Krümmungsradius der Hornhaut fokussiert. Dadurch wird die Wellenfront des einfallenden Strahls der Krümmung der Hornhaut angepasst und die Reflexion erfolgt für jeden Teilstrahl der Strahlung unter senkrechtem Winkel. Eine anschließende Parallelisierung der reflektierten Wellenfront erzeugt eine senkrecht zur optischen Achse stehende Wellenfront. Die Bewegung des Auges erzeugt eine globale Verkippung der reflektierten Wellenfront, da sich die Reflexionswinkel der Teilstrahlen in erster Näherung um einen gleichen Betrag ändern. Dieser Winkel wird durch die anschließende Parallelisierung der Linse nicht kompensiert. Die entstehende Verkippung kann so durch den Wellenfrontsensor-ASIC quantifiziert werden.

Abbildung 5.1 a) stellt die einfallende und reflektierte Wellenfront eines zentral zur optischen Achse positionierten Auges dar. Der Fokalfpunkt des einfallenden Lichtstrahls koinzidiert mit

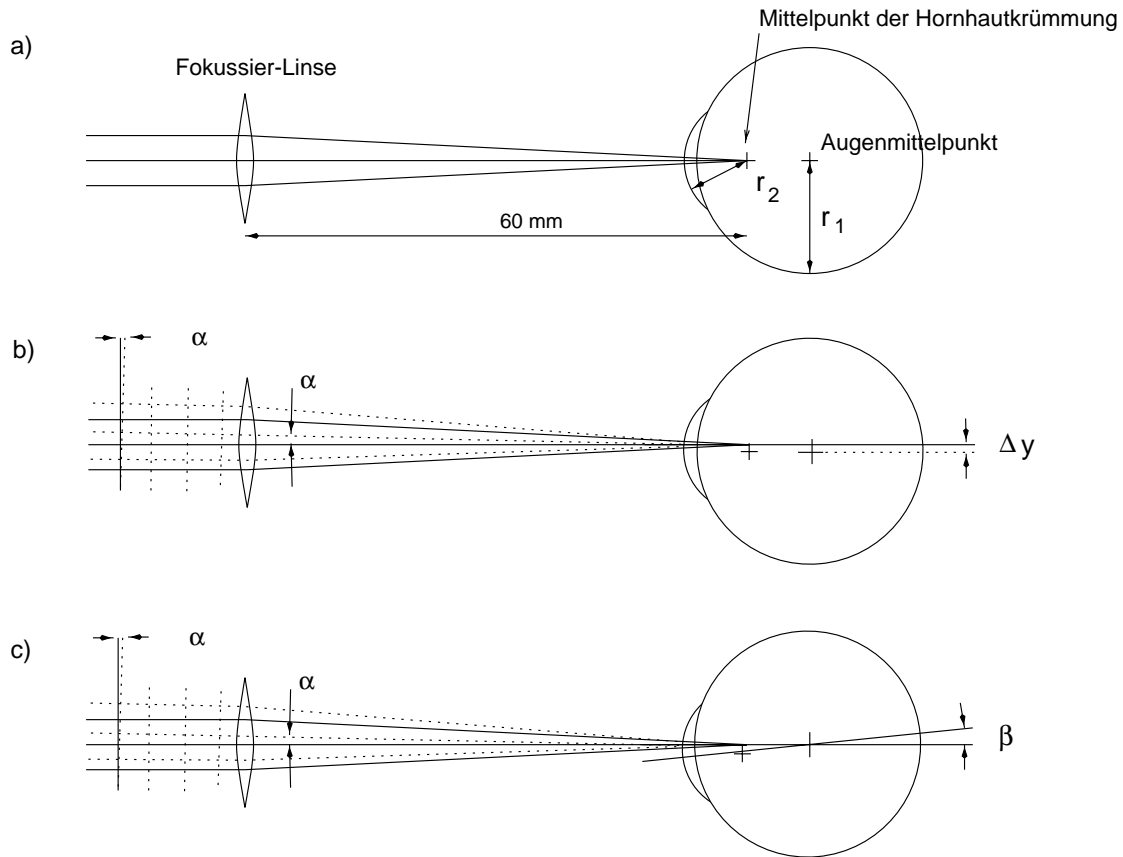


Abbildung 5.1: Wellenfrontverkipfung α bei rotatorischer und translatorischer Bewegung des Auges

dem Mittelpunkt des Krümmungsradius der Hornhaut. Die einfallende und die reflektierte Wellenfront besitzen die gleiche Form, so dass sich die Fokalfpunkte des Linsenarray in der Referenz-Position befinden.

In Teilbild b) ist die Eigenschaft der Reflexion unter Auftreten einer translatorischen Bewegung des Auges dargestellt. Die Mittelpunkte von Auge und Hornhautkrümmung verschieben sich gleichzeitig um denselben Betrag Δy aus der optischen Achse. Jeder Teilstrahl des einfallenden Lichtstrahls wird in erster Näherung unter einem gleichen Winkel reflektiert. Die Wellenfront des reflektierten Lichtstrahls ist global um einen Winkel α verkippt. Durch die Fokussier-Linse wird der reflektierte Lichtstrahl parallelisiert und eine globale Verkipfung der Wellenfront um den Winkel α folgt. Die Wellenfront trifft auf das Linsenarray des Wellenfront-sensors und entsprechend den in Abbildung 1.8 dargestellten Abbildungseigenschaften einer Linse verschieben sich die Fokalfpunkte des Linsenarray um Δx wie in Gleichung 1.7 angenähert. Der Wellenfrontsensor-ASIC detektiert die Verschiebung der Fokal-Punkte, gibt die Daten an das externe System weiter und der PC berechnet das Maß der Wellenfront-Verkipfung. Abbildung 5.2 stellt eine schematische Ansicht des translatorisch bewegten Auges dar. Unter Annahme der Fokallänge des Linsenarray von f_L und der Fokussierlinse von f_D berechnet sich die Verschiebung der Fokalfpunkte Δx zu

$$\Delta x = f_L \sin \alpha = \frac{f_L}{f_D} \Delta y \quad (5.1)$$

In Teilbild c) der Abbildung 5.1 ist die Stellung des Auge rotatorisch verändert, so dass sich der

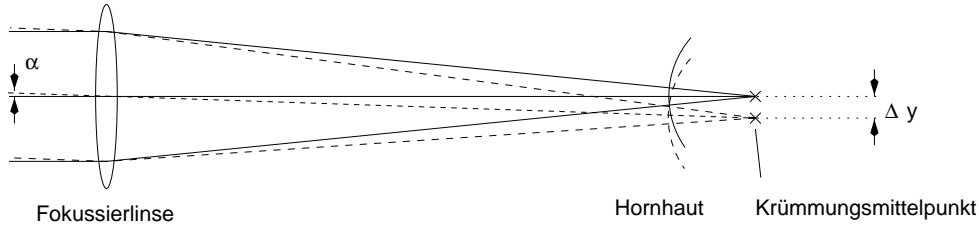


Abbildung 5.2: Berechnung des Winkels α bei translatorischer Bewegung. Der Maßstab der Auslenkung Δy ist aus Gründen der Darstellbarkeit vergrößert.

Krümmungsmittelpunkt der Hornhaut exzentrisch und der Augenmittelpunkt zentrisch zu der optischen Achse befinden. Es resultiert eine qualitativ gleichwertige Verkippung der reflektierten Wellenfront wie im Falle der translatorischen Bewegung des Auges die durch das Linsenarray in eine globale Verschiebung der Fokalfpunkte umgesetzt wird.

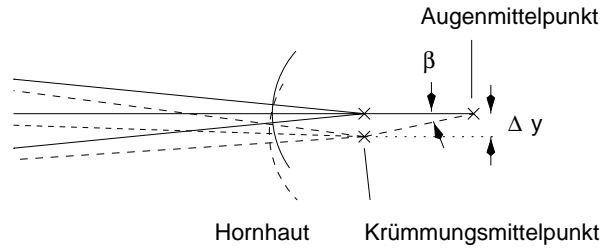


Abbildung 5.3: Berechnung einer rotatorischen Bewegung. Der Maßstab der Auslenkung Δy ist aus Gründen der Darstellbarkeit vergrößert.

Abbildung 5.3 stellt eine schematische Ansicht des rotatorisch bewegten Auges dar. Unter Annahme der gleichen Fokallänge des Linsenarray und der Fokussierlinse wie zuvor und des Radius des Auges r_1 und der Hornhautkrümmung r_2 berechnet sich Δx zu

$$\Delta x = f_L \sin \alpha = \frac{f_L}{f_D} \Delta y = \frac{f_L}{f_D} (r_1 - r_2) \sin \beta \quad (5.2)$$

Bei einer Rotation des Auges um $\beta = \pm 2.05^\circ$ oder einer translatorischen Bewegung des Auges um den Betrag $\Delta y = \pm 204 \mu m$ und unter Anwendung von $f_L = 53 mm$, $f_D = 53 mm$, $r_1 = 13.5 mm$ und $r_2 = 7.8 mm$ ergibt sich eine Abweichung der Fokalfpunkte von $\Delta x = 360 \mu m$. Dies stellt die Grenze des dynamischen Bereichs des verwendeten *Prototyp 1* dar, da die optisch aktive Fläche eines Clusters des ASIC eine Kantenlänge von ca. $360 \mu m$ besitzt. Die maximale Rotation, die in dem System detektierbar ist, entspricht daher einem Winkel von $\beta = 4.1^\circ$. Die maximale translatorische Bewegung entspricht einer Verschiebung von $\Delta y = 408 \mu m$.

Das Hartmann-Shack-Verfahren zur Messung einer Augenbewegung ist aufgrund der exakten Bestimmung der Verkippung einer reflektierten Wellenfront geeignet, kleinste Bewegungen des Auges zu detektieren. Unter Definition der Auflösung aufgrund einer minimal detektierbaren Abweichung eines Fokal-Punktes von $\Delta x = 45 \mu m$ resultiert nominal ein kleinster messbarer Winkel der Augenrotation von $\beta = 0.51^\circ$, entsprechend einer translatorischen Bewegung von $\Delta y = 50.9 \mu m$. Aufgrund der in praktischen Anwendungen auftretenden statistischen Verteilung der geringfügig unterschiedlichen subapertalen Wellenfrontverkippungen in den Linsen des Array

ist durch arithmetische Mittelung eine höhere Auflösung der globalen Verkipfung möglich. Bei einer globalen Verkipfung können die Fokalfpunkte mancher Cluster früher in den nächsten Detektions-Streifen (Diskretisierungs-Stufe) verschoben werden als andere und erzeugen eine Veränderung des gemittelten Signals.

5.3 Aufbau des Eye-Tracker

Das mechanische System, in das der Eye-Tracker eingebaut wurde, wurde im Rahmen von verschiedenen Projekten am Institut für Angewandte Physik der Universität Heidelberg entworfen. Es dient unter anderem dem Aufbau eines CCD-gestützten Wellenfrontsensors zur Vermessung der Aberrationen des Auges und wurde für den Einsatz des Wellenfrontsensor-ASIC als Eye-Tracker erweitert. Abbildung 5.4 zeigt die Anordnung der optischen und elektronischen Komponenten des Systems.

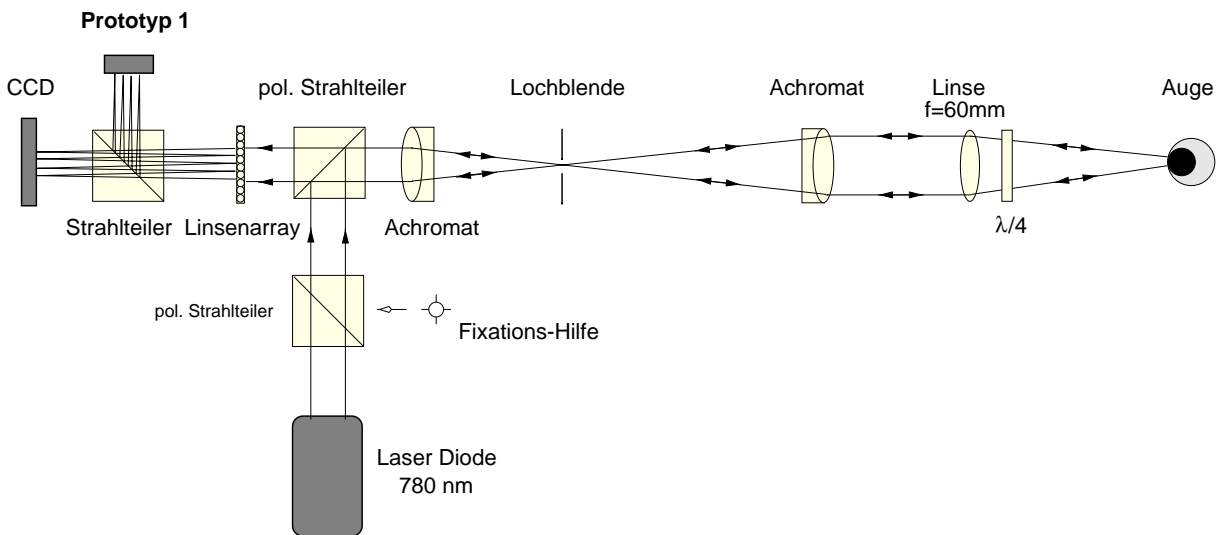


Abbildung 5.4: Der Aufbau des Eye-Tracker

Der Halbleiterdioden-Laser erzeugt einen kollimierten, linear polarisierten monochromatischen Lichtstrahl mit der Wellenlänge $\lambda = 780 \text{ nm}$ und einem Durchmesser von 8 mm. Der Strahl passiert einen Strahlteiler, an dem bei Bedarf ein optisches Muster zur Fixation des Auges in den Strahlengang eingeblendet werden kann, wird in dem nächsten Strahlteiler abgelenkt und durch ein Teleskop aus einem Achromaten-Paar, das eine Vergrößerung des Strahls aufgrund des Fokallängen-Verhältnisses von 60/40 besitzt, auf eine als Raumfilter eingesetzte Lochblende fokussiert. Anschließend erfolgt die Fokussierung des 12 mm breiten Strahls durch die Fokussier-Linse auf den Krümmungsmittelpunkt der Hornhaut und eine Änderung der Polarisations-Richtung auf rechts-zirkular durch ein $\lambda/4$ -Plättchen. Der an der Hornhaut spekulär reflektierte Strahl passiert dasselbe Plättchen, die Fokussier-Linse und das Teleskop. Aufgrund der rotierten Polarisationsrichtung durchdringt der Strahl den Strahlteiler, ohne abgelenkt zu werden und trifft auf das Linsenarray. In der Fokalebene des Linsenarray werden unter Verwendung eines Strahlteilers eine CCD-Kamera und der Wellenfrontsensor-ASIC beleuchtet, so dass die Fokalfpunkte durch beide Bildaufnahme-Systeme erfasst werden können.

Die CCD-Kamera ermöglicht die Kontrolle der von dem Linsenarray erzeugten Intensitäts-Verteilung der Fokalfpunkte und erlaubt die Justage des Systems. Da der Wellenfrontsensor-ASIC

einen geringen dynamischen Bereich der Fokalkpunkt-Detektion besitzt und daher sich die Fokalkpunkte vor Beginn eines Messzyklus' innerhalb der Cluster befinden müssen, erfolgt eine anfängliche Positionierung, um die Cluster des ASIC genau auf eine Referenz-Wellenfront einzustellen. Dies geschieht anhand des von der CCD-Kamera aufgenommenen Bildes. Da über den Strahlteiler die Fokalkpunkte gleichzeitig auf Wellfrontsensor-ASIC und die CDD-Kamera treffen, lässt sich ein Bereich innerhalb des CCD-Bildes definieren, in dem sich *virtuell* die Cluster des ASIC befinden. Die Justage des Systems erfolgt unter Positionieren der durch eine Referenz-Wellenfront erzeugten Fokal-Punkte innerhalb der virtuellen Cluster-Matrix des CCD-Bildes.

Das beschriebene Labor-System wurde an ein Gerät zur photorefraktiven Keratektomie (im Folgenden als Keratom bezeichnet) angebracht, um die Augenbewegungen von Probanden zu messen. Das verwendete Keratom befand sich im Auslieferungszustand und konnte für Probemessungen benutzt werden. Die veränderte Geometrie der Strahlführung erforderte eine Anpassung des Laborsystems, das als Fokussier-Linse eine Linse der Brennweite 60 mm verwendet. Im Keratom wurde als Fokussier-Linse die Linse zur Fokussierung des Excimer-Laserstrahls verwendet, die eine Fokallänge von 250 mm besitzt, so dass das Laborsystem um ein Teleskop durch eine zusätzliche 250 mm Linse erweitert werden musste. In Abbildung 5.5 ist die Strahlführung des Excimer-Laserstrahls des Keratoms und des Mess-Laserstrahls des Eye-Tracker eingezeichnet.

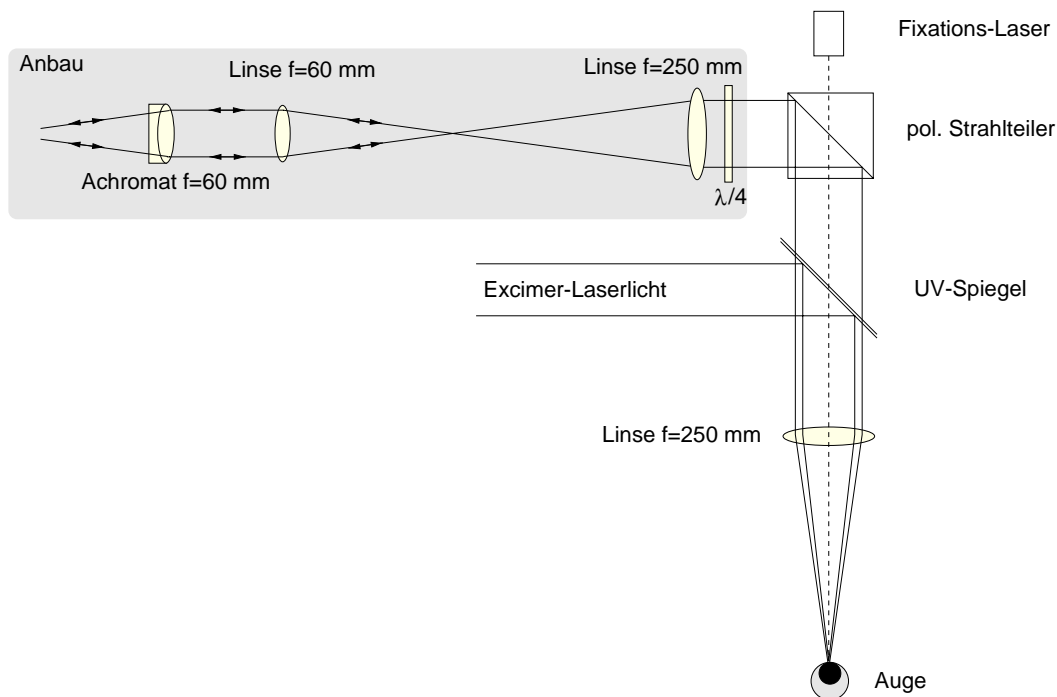


Abbildung 5.5: Der Anbau des Eye-Tracker an das Keratom

Der Mess-Laserstrahl, der aus der 60 mm-Fokussier-Linse des Labor-Systems austritt, trifft auf die erste Linse des zusätzlichen Teleskops und wird durch einen Strahlteiler des Keratoms, durch den im normalen Betrieb ein CCD-Beobachtungssystem eingespiegelt werden kann, in den Strahlengang eines Fixations-Laser eingebracht. Der Fixations-Laser ermöglicht dem Patienten die Kontrolle der Augenbewegung. Anschließend durchdringt der Mess-Laserstrahl einen UV-Spiegel, durch den der Ablations-Laserstrahl des Excimer-Lasers abgelenkt wird. Die letzte Linse vor dem Auge des Patienten fokussiert den Laser-Strahl des Excimer-Lasers auf die Hornhaut-Oberfläche und den Laserstrahl des Eye-Tracker auf den Krümmungsmittelpunkt der Hornhaut.

Der Mess-Laserstrahl wird reflektiert und verläuft entlang des beschriebenen Weges bis zu dem angebrachten Labor-System zurück. Dort erfolgt über die oben genannten Elemente die Generation der Fokalfpunkte durch das Linsenarray des Wellenfrontsensors und die Detektion der Fokalfpunkte-Verschiebungen durch den *Prototyp 1*.

5.4 Software des Eye-Tracker

Zur Verarbeitung der Daten des Eye-Tracker durch den PC wurde eine Software entwickelt, die über die Darstellung der Fokal-Punkte des Linsenarray und eine Berechnung der globalen Verkippung der Wellenfront eine Kontrolle der Augenbewegung ermöglicht. Das Auslesen des Wellenfrontsensor-ASIC erfolgt über das in Kapitel 4 beschriebene *System 1*, indem die generierten Daten des ASIC über den FPGA an den Parallel-Port des PCs übertragen werden. Abbildung

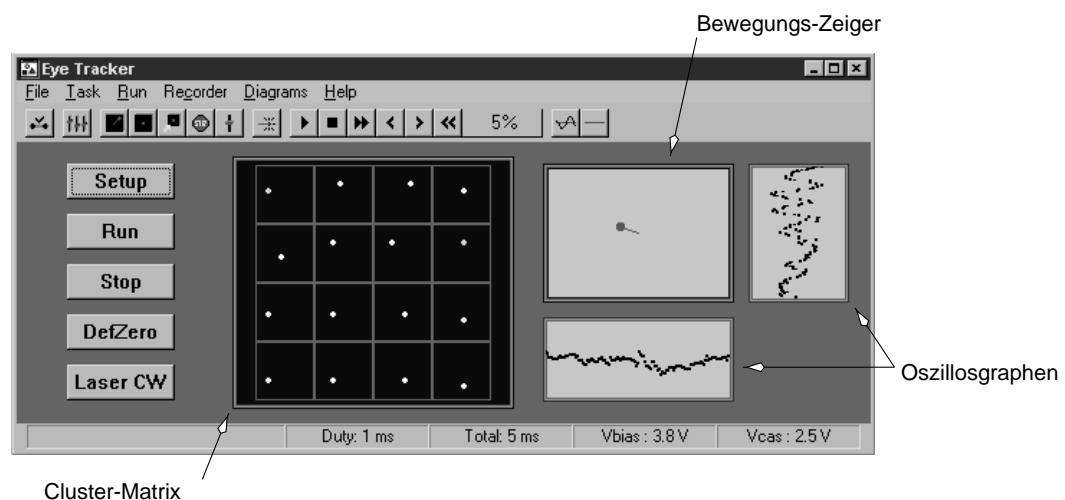


Abbildung 5.6: Die Eye-Tracker Software

5.6 stellt die Benutzeroberfläche der Software dar. Im Mittelpunkt der Oberfläche befindet sich eine Repräsentation der Roh-Daten des ASIC durch eine nachgebildete Cluster-Matrix. Die Fokalfpunkte der entsprechenden Cluster werden in Echtzeit dargestellt. Aufgrund der im Vergleich zur Aufnahmezeit langsameren Bilddarstellung des PCs wurde ein Burst-Modus der Messungen entwickelt, um die Echtzeit-Darstellung der Augenbewegung zu erlauben. Es erfolgten jeweils 8 Messungen in einer Frequenz von 1 kHz, in denen die Interrupt-Steuerung des Betriebssystems maskiert wurde, um die Fähigkeit des Systems zur Aufnahme der Messdaten mit einem zeitlichen Abstand von 1 ms zu demonstrieren. Anschließend wurden die zuletzt gemessenen Daten einer Sequenz nach Demaskierung der Interruptsteuerung dargestellt. In dem verwendeten System betrug die erforderliche Zeit zur Darstellung ca. 40 ms.

Rechts neben der Cluster-Matrix ist ein Zeiger definiert, der das Maß der Verkippung der Wellenfront aufgrund einer Umrechnung der Daten des ASIC darstellt. Die Positionsabweichungen der Fokalfpunkte des Clusters werden über alle Punkte arithmetisch gemittelt und ausgegeben. Neben dem Bewegungszeiger werden über zwei Oszilloskop-Felder zyklisch die vom Zeiger dargestellten letzten 1000 Werte für X- und Y-Position im zeitlichen Verlauf wiedergegeben.

Es besteht die Möglichkeit, die Werte der letzten 1000 Messwerte abzuspeichern und zu laden. Über eine Rekorder-Steuerung kann die Bewegung des Auges im transienten Verlauf wiederholt dargestellt werden.

5.5 Messungen der Augenbewegungen eines Probanden

Der beschriebene Eye-Tracker wurde zu Messzwecken an dem Keratom angebracht. Die Justage des Systems erfolgte unter Verwendung einer reflektierenden Kugel eines Durchmessers von ca. 15.6 mm, die in den Fokalkpunkt-Bereich der 250 mm-Linse des Keratoms gelegt wurde. Das angebaute System des Eye-Tracker wurde anschließend derart justiert, dass der Fokalkpunkt des Mess-Laserstrahls ca. 7.8 mm unterhalb des Fokalkpunktes des Excimer-Laserstrahls lag. Die erforderliche Justage des zusätzlichen Eye-Tracker-Teleskops erfolgte unter Verschieben der ersten 250 mm-Linse des Teleskops, ausgehend vom Eye-Tracker, und Zentrieren der optischen Achsen des Eye-Tracker und des Keratoms. Die optischen Elemente des Keratoms durften nicht zur Justage verwendet werden.

Die anfängliche Positionierung eines Probanden fand unter Benutzung von zwei sogenannten Pilot-Lasern des Keratoms statt, die sich genau im Fokus der 250 mm-Linse des Keratoms überschneiden. Der Kreuzpunkt dient als Anhaltspunkt zur Korrektur der vertikalen Position der Hornhautoberfläche. Anschließend wurde über die Verschiebetisch-Steuerung eine laterale Position des Probanden unter Kontrolle der visuellen Darstellung der Fokal-Punkte der nachgebildeten Cluster-Matrix der Eye-Tracker-Software angefahren und die Referenz-Wellenfront definiert. Der in Abbildung 5.5 eingezeichnete Fixations-Laser stellte ein optisches Ziel für das Auge dar.

Es zeigte sich, dass die Sensitivität des Eye-Tracker auf eine Bewegungen des Auges sehr hoch war und eine ausschließlich auf eine Rotation des Auges bezogene Messung der Verkipfung der Wellenfront nicht möglich war. Durch die instabile Anordnung der Patientenliege wurden translatorische Bewegungen des zu untersuchenden Auges verursacht, die den dynamischen Messbereich des Eye-Tracker überstiegen und die Aufnahme durch den Sensor vereitelten. Um die Funktionsfähigkeit des Laborsystems an dieser Stelle hervorzuheben, wird eine von dem in Abbildung 5.4 dargestellten System aufgenommene Messreihe vorgestellt.

Mit Hilfe einer Kopfstütze war es möglich, die Bewegungen des Kopfes innerhalb des dynamischen Bereichs zu fixieren. Es wurden mehrere Messreihen von Augenbewegungen aufgenommen, in denen der Proband eine oszillierende Bewegung der Augen vollführte.

Abbildung 5.7 stellt die durch den Eye-Tracker aufgenommene Messreihe der Augenbewegung eines Probanden dar. Die Sequenz besitzt eine zeitliche Länge von 6 Sekunden und gibt eine oszillierende Augenbewegung des Probanden wieder. Die Oszillation fand mit einer Frequenz von ca. 1 Hz statt und besaß eine Amplitude von rotatorisch ca. 2.4° oder translatorisch ca. $240 \mu\text{m}$. Die Bewegung in y-Richtung ist gegenüber der Bewegung in x-Richtung verringert und besitzt eine maximale Amplitude von rotatorisch ca. 1.3° und translatorisch ca. $130 \mu\text{m}$.

Die Messreihen zeigten, dass das System des Eye-Tracker eine sehr hohe Empfindlichkeit in der Detektion von Augenbewegungen besitzt. Durch den Einsatz des *Prototyp 1* des Wellenfrontsensor-ASIC wurde eine nominale Auflösung der Augenbewegung von rotatorisch ca. 0.5° ermöglicht. Der dynamische Bereich beträgt ca. 4° . Eine Erhöhung des dynamischen Bereichs des Eye-Tracker ist unter Einsatz eines Linsenarrays mit einer kürzeren Fokallänge möglich. Eine Fokallänge von zum Beispiel 27 mm könnte den dynamischen Bereich der translatorischen Bewegungsdetektion auf ca. $\pm 400 \mu\text{m}$ vergrößern und den dynamischen Bereich der rotatorischen Bewegungsdetektion auf ca. $\pm 4^\circ$.

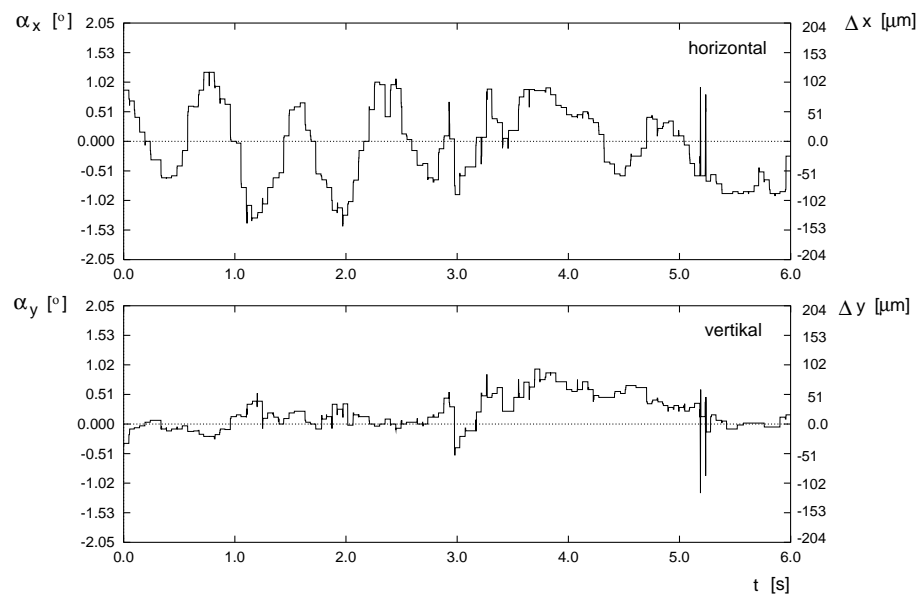


Abbildung 5.7: Gemessene Augenbewegung eines Probanden durch das Laborsystem

Kapitel 6

Messergebnisse des HSSX

Die in Kapitel 3 beschriebene Entwicklung des *HSSX* erfolgte unter Betracht der in Abschnitt 1.6 genannten Spezifikationen. Das System der aktiven Aberrationskompensation zur Auflösungsverbesserung des Retina Scan Tomographen erfordert eine hohe Signalverarbeitungsgeschwindigkeit in der Bestimmung der Wellenfront von bis zu 1 kHz bei einer reflektierten Photonenintensität von $J_{FP} \approx 200 \frac{pW}{Spot}$. In diesem Kapitel werden die Messergebnisse des *HSSX* vorgestellt und das Erreichen der Spezifikationen bewertet.

Zu Beginn erfolgt basierend auf Messergebnissen eine Abschätzung der in der Herstellung des ASIC mit dem 0.6- μ m-CMOS-Prozess der Firma *AMS* erreichten Übereinstimmung der Transistor-Charakteristika von identischen Transistor-Strukturen. Die Parameter-Streuung von verschiedenen Chips und die Qualität der Übereinstimmung von speziellen Test-Strukturen desselben Chips werden quantifiziert. Die Messung der Genauigkeit der Fokalkpunkt-Detektion unter Verwendung eines einzelnen, in der Größe und Intensität justierbaren Fokalkpunktes geben im weiteren Verlauf des Kapitels einen Einblick in die tatsächlich erreichbare Genauigkeit der Positionsdetektion. Die Messungen werden auf die Gestalt eines Fokalkpunktes des Linsenarray erweitert und das durch die Verwendung der Rückkopplung entstehende Rauschen der Positions-Detektion gemessen. Anschließend werden Messergebnisse optischer Aberrationen unter Verwendung eines Linsenarray dargestellt. Ein System wurde aufgebaut, mit dem die Aberrationen von optischen Medien bei Fokalkpunkt-Lichtleistungen von ca. 1 bis 2 nW und Integrationszeiten von 1 ms mit einer absoluten Genauigkeit von ± 0.16 dpt vermessen werden konnten. Den Abschluss des Kapitels bilden ein Überblick über die gemessene elektrische Verlustleistung des *HSSX* und eine Diskussion aller Messergebnisse.

6.1 Parameter des 0.6- μ m-CMOS-Prozess

Wichtigstes Merkmal der Prozessqualität eines ASIC der analogen Signalverarbeitung ist die Übereinstimmung von charakteristischen Eigenschaften der MOS-FETs. Während in der digitalen Signalverarbeitung das Übereinstimmen der Transistoreigenschaften aufgrund der Diskretisierung von Spannungspegeln in zwei gegensätzliche Werte nicht kritisch ist, besteht in Systemen mit identischen Strukturen einer analogen optischen Signalverarbeitung ein Interesse an einem über der Chip-Fläche homogenen Verhalten der Bauelemente.

Die Ausbeute in der Herstellung eines ASIC in der 0.6- μm -CMOS-Technologie der Firma *AMS* wird von *AMS* mit 40% beziffert. Diese Angabe bezieht sich auf die Berechnung der Wahrscheinlichkeit, mit der Wellenfrontsensor-ASIC keinen durch die Herstellung verursachten Defekt besitzen. Die Angabe bezieht sich nicht auf die Qualität der Funktionalität des ASIC selbst, da eine Abschätzung der Anzahl der im Rahmen der Spezifikationen funktionierenden Chips erst nach Untersuchung der Transistor-Charakteristika und Messung der Funktionalität selbst möglich ist.

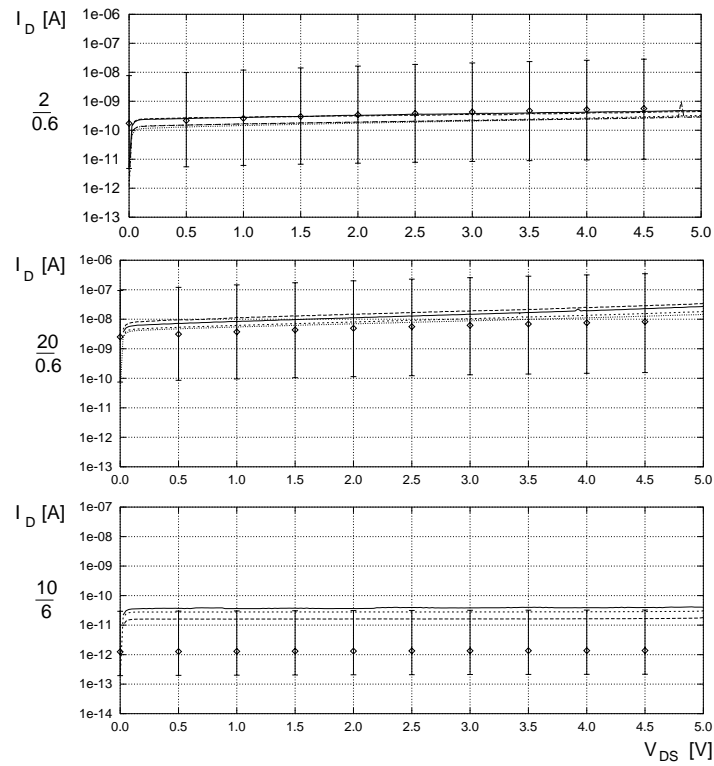
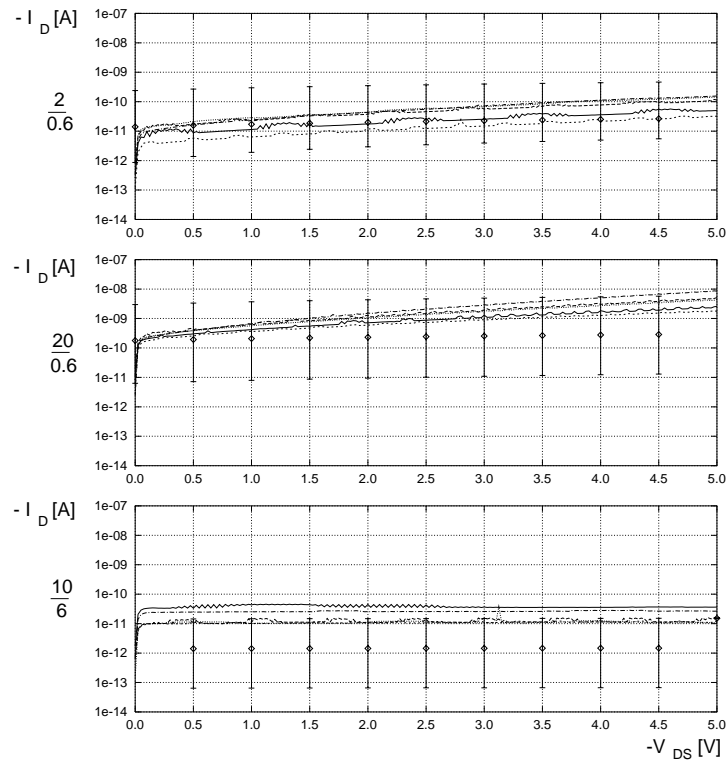
Ein Herstellungsprozess wird durch die Verteilungen der Parameter spezifiziert und die Simulation der Schaltungen erfolgt unter Verwendung der in Kapitel 3 genannten Eck-Parameter, die die jeweils äußeren Bereiche dieser Verteilungen festlegen. Die Eck-Parameter umfassen die Eigenschaften aller in dem Prozess hergestellten Strukturen und modellieren zum Teil mehrere Dekaden umfassende Unterschiede der Drain-Ströme von Transistoren unter gleichen Spannungen. Während die Streuung der Parameter zwischen den verschiedenen Chips von verschiedenen Prozessdurchläufen durchaus die durch die Eck-Parameter festgelegte Verteilung überdecken kann, ist die Übereinstimmung von auf derselben Chip-Oberfläche befindlichen Elementen im Allgemeinen größer. Diese Übereinstimmung ist in der analogen, optischen Signalverarbeitung der Vision-Chips notwendig, da die vielfach wiederholten Schaltungen möglichst ähnliche Funktionalitäten besitzen sollen, um identische Übertragungsfunktionen zu realisieren. Aufgrund der Unterschiede der Transistoreigenschaften benachbarter Strukturen entstehen in Vision-Chips jedoch Differenzen, die aufgrund der festen Lokalität als *Fixed-Pattern-Noise* bezeichnet werden.

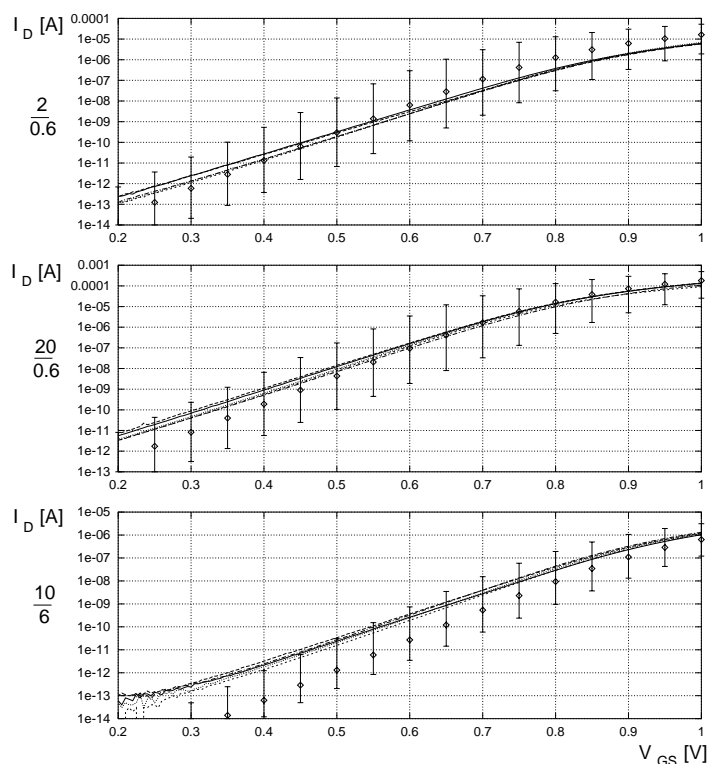
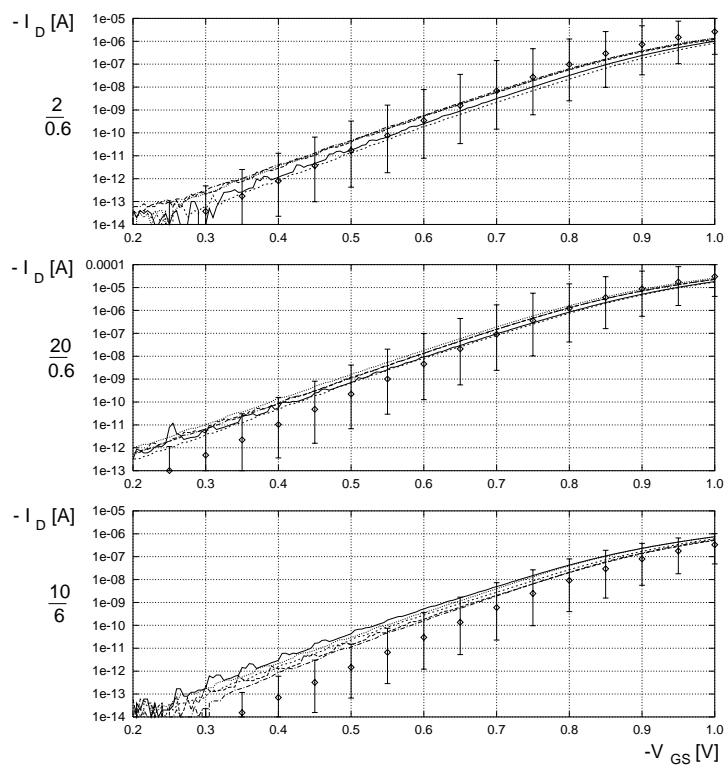
Im Folgenden werden die Messergebnisse von MOS-FET-Teststrukturen des *HSSX* vorgestellt und mit den Eck-Parametern der Prozess-Spezifikationen verglichen. Die Bewertung der absoluten Übereinstimmung der Drain-Ströme von speziellen, benachbarten Transistor-Strukturen unter Anlegen identischer Spannungen folgt anschließend. Insgesamt wurden MOS-FET-Kennlinien von 32 Chips aufgenommen. 5 Chips wurden für die Bewertung der Unterschiedlichkeit der MOS-FET-Charakteristika zwischen einzelnen Chips und 32 Chips zur Bewertung der Unterschiedlichkeit benachbarter identischer Strukturen desselben Chips vermessen. Aufgrund der Erkenntnisse der Übereinstimmungs-Qualität der untersuchten Chips wurden Exemplare ausgewählt, mit denen ein System zur Bestimmung optischer Aberrationen aufgebaut wurde.

Übereinstimmung der Transistor-Kennlinien verschiedener Chips

Die Messung der charakteristischen Kennlinien von MOS-FET erfolgte unter Auswahl von fünf Chips. Jeder Chip verfügt über insgesamt 16 Test-MOS-FET, deren Anschlüsse über Probe-Pads für Messungen zugänglich sind. Es handelt sich um drei verschiedene Transistor-Geometrien mit jeweils gegensätzlichem Leitfähigkeits-Typ. Die W/L-Verhältnisse betragen 2/0.6, 20/0.6 und 10/6. Die beiden ersten Geometrien entstammen Test-Strukturen der Firma *AMS*, die letzte einem eigenen Entwurf. Aus Messung der Charakteristika dieser verschiedenen NMOS-FET und PMOS-FET kann die Qualität der zur Schaltungssimulation benutzten Modelle abgeschätzt und die Einstellung der für die Funktionalität erforderlichen Bias-Spannungen berechnet werden.

In Abbildung 6.1 und 6.2 sind die Ausgangs-Kennlinien $I_D = f(V_{DS})$ der 6 verschiedenen NMOS- oder PMOS-FET von 5 Chips im Subthreshold-Bereich dargestellt. Zusätzlich ist die Toleranz der Simulationen unter Verwendung des *AMS*-Modells *Ams15* abgebildet. Die mittleren Werte der Fehlerbalken repräsentieren den Modellsatz der typischen Parameter tm , die unteren und die oberen Werte repräsentieren die Kennlinien unter Auswahl der Eck-Parameter ws und wp , entsprechend *schwacher* oder *starker* MOS-FET. Die Gate-Source-Spannung wurde auf $|V_{GS}| = 500\text{ mV}$ eingestellt und die Drain-Source-Spannung von 5 V bis 0 V, respektive -5 V bis 0 V verändert.

Abbildung 6.1: I_D gegen V_{DS} identischer NMOS-FET-Strukturen verschiedener ChipsAbbildung 6.2: I_D gegen V_{DS} identischer PMOS-FET-Strukturen verschiedener Chips

Abbildung 6.3: I_D gegen V_{GS} identischer NMOS-FET-Strukturen verschiedener ChipsAbbildung 6.4: I_D gegen V_{GS} identischer PMOS-FET-Strukturen verschiedener Chips

Es zeigt sich, dass die Kennlinien der AMS-Strukturen durch die Simulations-Kennlinie sehr gut angenähert werden. Diese Tatsache ist nicht überraschend, da die Erstellung der Parametersätze des Modells *Ams15* durch die Firma AMS genau aufgrund dieser Transistor-Strukturen geschieht. Die Kennlinie der selbst definierten Struktur mit $W/L=10/6$ wird durch das Modell nicht gut beschrieben. Die Simulation mit den *tm*-Parametern erzielt einen um den Faktor von ca. 15 geringeren Drain-Strom.

Die Steigung der Kennlinien, die ein Maß für den Ausgangswiderstand des MOS-FET ist, wird für alle drei Strukturen gut angenähert. Die Berechnungen und Simulationen der Ausgangsbufferspannungen der WTA-Zellen und der Inverter-Querströme der Schieberegisterzellen sind daher für realistische Bereiche vorgenommen worden.

In Abbildung 6.3 und 6.4 sind die Übertragungs-Kennlinien $I_D = f(V_{GS})$ der NMOS- oder PMOS-FET im Subthreshold-Bereich bei $|V_{DS}| = 2V$ dargestellt. Die Qualität der Simulation der AMS-Strukturen ist sehr gut, die Kennlinie der selbst definierten Struktur ist zufriedenstellend. Die Steigung der Kennlinien im exponentiellen Bereich ist in der Simulation stärker als die praktischen Messungen zeigen. Der Übertragungsleitwert nach Gleichung 2.21 ist abhängig von dem die Steigung des exponentiellen Verlaufs bestimmenden Faktor n und wird in der Simulation überhöht modelliert. Die Bias-Spannungen zur Erzeugung der benötigten Ruhestrome der einzelnen Schaltungsmodule müssen daher gegenüber den in der Simulation verwendeten Werten erhöht werden.

Unter Bewertung der Kennlinien lässt sich die Abweichung der Drain-Ströme der identischen Strukturen von verschiedenen Chips berechnen. Von den 5 Messwerten der Drain-Ströme der identischen Strukturen an jedem Messpunkt der entsprechenden Kennlinie wurde die Standardabweichung berechnet. Der Wert σ stellt die auf den Durchschnitt der 5 Werte normierte Standardabweichung der Werte dar. In Tabelle 6.1 ist der jeweils größte und kleinste Wert σ_{min} und σ_{max} der entsprechenden Kennlinie der 5 Transistoren eines Typs eingetragen. Der Wert $\bar{\sigma}$ bildet den Durchschnitt aller σ einer Kennlinie. Zur Berechnung der Werte der Übertragungskennlinien $I_D = f(V_{GS})$ wurden die Drain-Ströme bei Gate-Source-Spannungen $|V_{GS}| < 400mV$ nicht mehr berücksichtigt, da in diesem tiefen Subthreshold-Bereich die Drain-Ströme Werte annehmen, die durch den verwendeten Parameter-Analysator nicht mehr verlässlich erfasst werden konnten. Aus der Tabelle kann entnommen werden, dass die Übereinstimmung von PMOS-FET verschie-

Kennlinie Typ	$I_D = f(V_{DS})$ NMOS			$I_D = f(V_{DS})$ PMOS			$I_D = f(V_{GS})$ NMOS			$I_D = f(V_{GS})$ PMOS		
Geometrie	$\frac{2}{0.6}$	$\frac{20}{0.6}$	$\frac{10}{6}$	$\frac{2}{0.6}$	$\frac{20}{0.6}$	$\frac{10}{6}$	$\frac{2}{0.6}$	$\frac{20}{0.6}$	$\frac{10}{6}$	$\frac{2}{0.6}$	$\frac{20}{0.6}$	$\frac{10}{6}$
σ_{min}	0.24	0.29	0.34	0.45	0.16	0.52	0.07	0.15	0.10	0.20	0.17	0.17
σ_{max}	0.35	0.37	0.39	0.58	0.60	0.74	0.36	0.37	0.33	0.70	0.44	0.82
$\bar{\sigma}$	0.30	0.34	0.36	0.51	0.39	0.62	0.19	0.23	0.23	0.42	0.32	0.44

Tabelle 6.1: Auf den Mittelwert normierte Standardabweichungen der I_D von 5 Chips

der Chips schlechter ist als die der NMOS-FET. Besonders im tiefen Subthreshold-Bereich der PMOS-FET beträgt σ_{max} 70-80%. Die Übereinstimmung breiter PMOS-FET-Strukturen mit großem W ist signifikant besser als die der schmalen Strukturen. Die Simulation der Schaltung unter Wahl des *Ams15*-Modells ist für die typischen Transistor-Geometrien mit kurzem Kanal hinreichend genau, für das Verhalten anderer Transistor-Geometrien, in denen das Verhältnis aus Gate-Breite und -Länge kleiner 1 ist, erfolgt eine schlechte Modellierung. Generell wird der

Faktor n des Übertragungsleitwertes der MOS-FET im Subthreshold-Bereich durch die Simulation überhöht, die tatsächliche Transkonduktanz ist entsprechend geringer anzunehmen.

Übereinstimmung benachbarter identischer Strukturen

Die Übereinstimmung der benachbarten Strukturen ist für die Funktionalität des *HSSX* wichtig, da aufgrund der begrenzten Fläche der Signalverarbeitung zwischen den optisch aktiven Bereichen der Cluster keine Funktion zur Selbstkalibration angeordnet werden konnte, um die Konsequenzen von Mismatching zu reduzieren. Um die Qualität der Übereinstimmung benachbarter Strukturen abschätzen und Chips mit potentiell geringem Mismatching aussuchen zu können, wurden die Übertragungskennlinien von selbstdefinierten Teststrukturen vermessen. Es handelt sich um die MOS-FETs mit einem W/L -Verhältnis von 10/6, die auch schon bei den Chipübergreifenden Messungen benutzt wurden. In Abbildung 6.5 ist das Layout der Test-Struktur wiedergegeben.

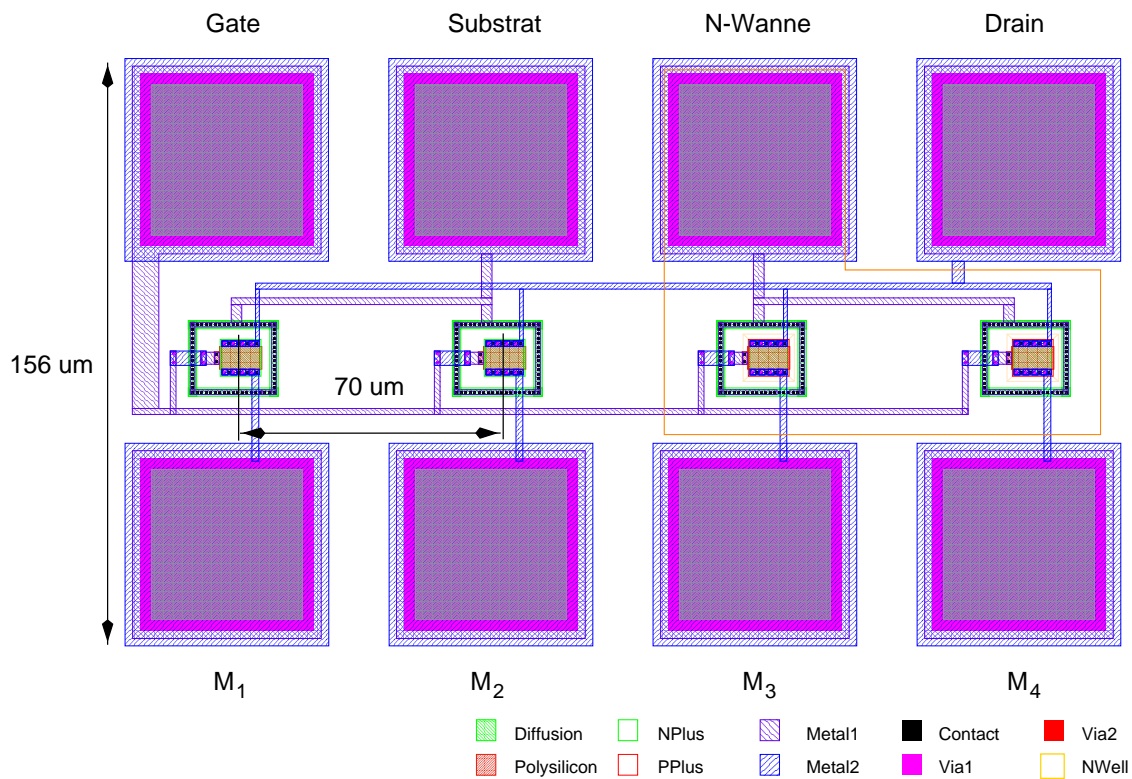


Abbildung 6.5: Teststruktur zur Messung von Mismatching

Jeweils zwei identische Strukturen des p- und des n-Typs sind platziert und durch einen geschlossenen Substratkontakt-Ring mit einem definierten Substrat- oder N-Wannen-Potential versorgt. Der Abstand eines Paares beträgt $70 \mu\text{m}$. Die Test-Struktur ist in einem Abstand von $520 \mu\text{m}$ wiederholt. Die Anordnung innerhalb des Layout des *HSSX* ist in Abbildung 3.32 von Kapitel 3 zu sehen.

Um das Übereinstimmen der identischen Strukturen im Abstand von $70 \mu\text{m}$ und $520 \mu\text{m}$ zu quantifizieren, wurden jeweils drei Kennlinien-Paare desselben Transistor-Typs vermessen. Dabei wurden die Übertragungs-Kennlinien der im Abstand von $70 \mu\text{m}$ befindlichen Strukturen in ein Verhältnis gesetzt, sowie die der beiden jeweils äußeren Strukturen der in den im Abstand von $520 \mu\text{m}$ angeordneten Test-Layouts. Für die beiden Leitfähigkeitstypen ergeben sich daraus

insgesamt 6 Kennlinne-Paare, die mit den Buchstaben A bis F gekennzeichnet werden, siehe Abbildung 6.6.

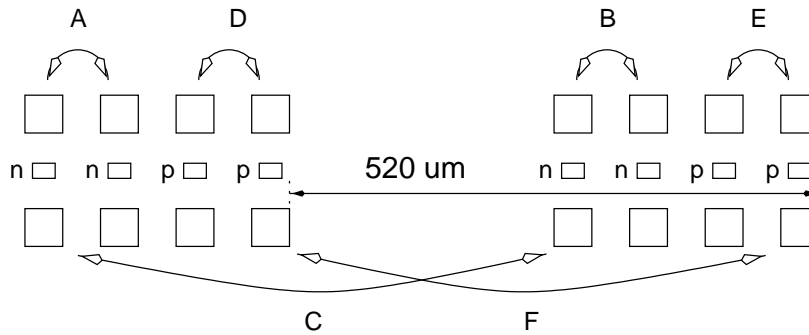


Abbildung 6.6: Vermessene Paare

Die Messung der Kennlinien erfolgte mit einer Drain-Source-Spannung von $|V_{DS}| = 2\text{ V}$ und einer variierten Gate-Source-Spannung zwischen 1 V und 200 mV, respektive -1 V und -200 mV. Um speziell die Übereinstimmung der PMOS-FET unter Ansteuerung des Bulk zu vermessen, wurde ein Bulk-Source-Potential der PMOS-FET von 500 mV eingestellt¹.

In Tabelle 6.2 sind die Ergebnisse der Messungen der 32 Chips aufgelistet. Die Werte stellen die auf den jeweils ersten Transistor des Paares normierte maximale Differenz der beiden Drain-Ströme dar. Dazu wurden die einzelnen Werte der beiden Übertragungs-Kennlinien eines Paares subtrahiert, das Ergebnis auf den Betrag des Stromes des ersten Transistors normiert und der maximale Wert in einem Spannungsbereich von $600\text{ mV} < V_{GS} < 1\text{ V}$ bzw. $-1\text{ V} < V_{GS} < -600\text{ mV}$ in die Tabelle eingetragen. In der letzten Spalte der Tabelle ist die Standardabweichung σ_1 der Übereinstimmungswerte aller 6 Paare eines Chips eingetragen. In der letzten Zeile ist die Standardabweichung σ_2 der Übereinstimmung der Paare A bis F aller vermessenen Chips berechnet.

Aus der Tabelle und den Ergebnissen der Messungen von Strukturen verschiedener Chips lässt sich folgern, dass zwischen benachbarten Strukturen eine bessere Übereinstimmung besteht als zwischen identischen Strukturen verschiedener Chips. Die Unterschiedlichkeit der PMOS-FET ist stärker, da die Transistoren, wie in Tabelle 6.1 bei den Ausgangskennlinien von n- und p-Typ zu sehen, generell über ein größeres Mismatching verfügen und außerdem über die Ansteuerung des Bulk in einen tieferen Subthreshold-Bereich geführt werden, in dem sich die Übereinstimmung stark reduziert. Aufgrund der geringeren Standardabweichungen σ_2 der Paare A und B gegenüber C, sowie D und E gegenüber F ist es möglich, eine bessere Übereinstimmung der im Abstand von $70\text{ }\mu\text{m}$ angeordneten Paare als die der im Abstand $520\text{ }\mu\text{m}$ benachbarten Paare abzuschätzen. Aufgrund einer geringen Standardabweichung σ_1 sind Chips für die weiteren Messungen ausgewählt worden.

Die Messungen der Übereinstimmung der Transistoreigenschaften zeigen, dass die Problematik des Mismatching durch Anwendung eines identischen Layout der Elemente im Allgemeinen etwas reduziert werden kann. Die Tabelle 6.2 zeigt jedoch, dass es selbst zwischen benachbarten identischen Strukturen teilweise zu einer Abweichung der Drain-Ströme bei gleichen Anschluss-Spannungen von ca. 21% kommen kann, siehe Chip H8. Generell lässt jedoch ein geringer Wert der Abweichung zweier benachbarter Strukturen auf eine lokale homogene Qualität des Prozesses während der Herstellung der beiden Strukturen schließen, die sich über der lateralen

¹Dies geschah im Hinblick auf die Verwendung des Body-Effektes zur Einstellung eines willkürlichen Spiegel-Verhältnisses des Rückkopplungs-Pfades der WTA-Zellen.

Chip	Paar						σ_1
	A	B	C	D	E	F	
H3	0.01	-0.01	-0.01	0.04	0.06	-0.02	0.03
H4	0.01	0.00	0.00	-0.02	0.04	-0.02	0.02
H5	0.02	-0.04	0.00	-0.20	0.04	-0.20	0.11
H6	0.01	-0.01	0.02	0.01	0.08	-0.10	0.06
H7	0.00	-0.05	0.03	-0.12	0.11	0.06	0.08
H8	0.02	0.12	-0.40	0.12	0.13	0.13	0.21
H9	-0.01	-0.04	0.04	0.04	0.05	-0.18	0.09
H10	0.04	0.03	-0.06	-0.09	0.12	-0.08	0.08
H11	0.02	-0.03	0.01	0.05	0.02	0.10	0.04
H12	0.09	0.04	0.11	-0.14	-0.18	-0.06	0.12
H14	0.05	0.02	0.03	-0.04	-0.04	-0.02	0.04
H15	0.04	0.02	-0.06	-0.09	0.00	-0.06	0.05
H16	0.03	-0.25	-0.02	0.12	0.10	0.05	0.13
I1	-0.02	-0.04	-0.05	-0.09	-0.05	-0.02	0.03
I2	0.02	0.04	0.02	-0.09	-0.20	0.20	0.13
I3	0.02	-0.02	-0.04	-0.02	-0.04	-0.14	0.05
I4	0.02	-0.02	-0.02	0.02	-0.08	0.08	0.05
I5	-0.02	0.02	-0.25	-0.02	-0.04	-0.10	0.10
I7	-0.04	0.07	-0.04	-0.02	-0.01	-0.05	0.04
I8	0.02	-0.02	-0.02	0.02	0.06	0.04	0.03
I9	0.01	0.12	-0.09	0.02	0.02	0.01	0.07
I10	0.04	0.02	0.03	0.05	0.02	-0.05	0.04
I12	0.06	0.04	-0.01	-0.04	-0.05	-0.09	0.06
I13	-0.02	0.02	-0.04	0.05	0.14	-0.08	0.08
I14	0.03	0.08	0.01	0.09	-0.02	0.07	0.04
I15	-0.02	0.01	0.01	0.18	-0.04	0.16	0.10
I16	-0.02	-0.01	0.01	0.11	0.06	0.07	0.05
A1	0.02	-0.01	0.03	0.06	0.10	0.07	0.04
A5	-0.02	0.01	0.02	0.16	-0.02	0.20	0.10
B16	-0.06	-0.02	-0.03	0.18	-0.30	0.25	0.19
C11	-0.05	0.03	-0.06	-0.09	-0.07	0.15	0.09
H11	0.02	-0.03	0.01	-0.04	-0.02	-0.04	0.03
σ_2	0.03	0.06	0.09	0.09	0.10	0.11	

Tabelle 6.2: Normierter maximaler Unterschied der Drain-Ströme in Übertragungskennlinien zweier identischer, benachbarter Transistor-Strukturen

angrenzenden Fläche erstrecken kann. Insofern ist die Auswahl von Chips mit einem geringen σ_1 sinnvoll, garantiert jedoch nicht zwangsläufig eine große Homogenität der Chips.

Die Ergebnisse der Messungen dieses Abschnitts zeigen, dass die Übereinstimmungen von identischen benachbarten Strukturen großen Schwankungen unterliegen, die sich in der Entwicklung der Schaltung eines ASIC nicht durch die Anwendung der typischen Simulations-Modelle abschätzen lassen. Zur Verringerung dieser Ungewissheit können durch besondere Layout-Gestaltungen relevante Schaltungsteile derart angeordnet werden, dass vor allem die lateralen Gradienten der im Prozess verwendeten Diffusions-, Implantations- oder Epitaxie-Schritte

oder die im Betrieb auftretenden Temperaturgradienten ein nur geringes Mismatching verursachen. Die typischen Anordnungen hierfür sind Verschachtelungen der entsprechenden MOS-FET-Strukturen untereinander. Diese Layout-Möglichkeiten sind jedoch flächenfordernd oder aufgrund der über weite Bereiche der Chip-Oberfläche verteilten Elemente nicht möglich.

In der Entwicklung des *HSSX* bestand eine Restriktion in der zur Signalverarbeitung vorhandenen Fläche zwischen den optisch aktiven Bereichen der Cluster. Um einen sinnvollen dynamischen Bereich der Fokuspunkt-Positionsdetektion zu ermöglichen, wurden die Elemente der Signalverarbeitung äußerst dicht angeordnet, so dass keine Möglichkeit der speziellen Verschachtelung der auf Mismatching sensitiv reagierenden Elemente bestand. Daher wurden durch die vorangehenden Messungen Chips ausgewählt, deren Mismatching per se gering war.

In den folgenden Abschnitten wird die Homogenität der Cluster-Matrix und der Cluster des *HSSX* behandelt.

6.2 Messung der Homogenität der Cluster-Matrix

Der *HSSX* verfügt über insgesamt 6 Steuer-Spannungen, die zur Funktionalität der WTA-Module auf bestimmte Werte eingestellt werden müssen. Tabelle 6.3 stellt die Spannungen, sowie ihre Funktion und im Falle von Bias-Spannungen die abgeschätzten erzeugten Bias-Ströme der Transistoren dar.

Name	Funktion	typischer Wert [V]	I_D
V_{bout}	Bias-Spannung des Ausgangsbuffers	0.74	30 nA
V_{cas}	Bias-Spannung der regulierten Kaskode	0.75	35 nA
V_{dmy}	Bias-Spannung zur Erzeugung des Dummy-Stromes	4.4	100 pA
V_{src}	Bias-Spannung zur Erzeugung des Stromes I_{src}	0.82	60 nA
V_{init}	Spannung zu Initialisierung der Knoten Ψ	0.7	
V_{bulk}	Bulk-Spannung der Feedback-Transistoren	6.2	

Tabelle 6.3: Bias-Spannungen und abgeschätzte resultierende Bias-Ströme einer WTA-Zelle

Die Wahl der Spannungen folgt den Überlegungen zur Funktionalität der Schaltung aus Kapitel 3. Nach den Ausführungen des vorangehenden Abschnitts der Prozessparameter-Bestimmung unterliegen die Drain-Ströme der Bias-Transistoren bei Anlegen von identischen Spannungen gewissen Schwankungen, die unterschiedliche Arbeitspunkte jeder einzelnen Zelle verursachen. Nach den Ausführungen der vorangehenden Abschnitte kann der Unterschied einige 10% betragen. Die Inhomogenität der Cluster-Matrix des *HSSX* wurde in Bezug auf die oben genannten Bias-Spannungen quantifiziert.

Um eine Aussage über das Mismatching der Cluster zu machen, wurden die Bias- und Steuer-Spannungen gemäss Tabelle 6.2 gewählt und die Spannung V_{dmy} derart eingestellt, dass alle Dark-Bits der X- und Y-Richtungen der Cluster gesetzt waren. In diesem Zustand konnten die Potentiale der Knoten Ψ der entsprechenden WTA-Zellen² nicht genügend ansteigen, um Ströme in den Flow-Transistoren zu erzeugen, die in Antwort-Signale hätten umgesetzt werden können. Der Versuch wurde unter Abdecken der Photodetektor-Matrix vorgenommen, um den Einfluss von Photonen zu verhindern.

²Abbildung 3.24 von Kapitel 3 stellt eine WTA-Zelle dar

Anschließend wurde V_{dmy} stetig verringert, so dass der Stromfluss der Dummy-Zellen erhöht wurde. Ab einer bestimmten Spannung kann die Dummy-Zelle Ausgangssignale erzeugen, da die Potentiale der Knoten Ψ der Dummy-Zellen Werte erreichen, die einen gegenüber den restlichen M_F der anderen Zellen der WTA-Module erhöhten Stromfluss verursachen. Aufgrund des Mismatching aller in den WTA-Zellen befindlichen Transistoren resultierte ein Spannungsbereich von V_{dmy} , innerhalb dessen die verschiedenen Cluster nach und nach das Ausgangssignal generieren und somit das entsprechende Dark-Bit löschen.

In Abbildung 6.7 und 6.8 ist der gemessene Spannungsbereich von V_{dmy} für die X- und die Y-WTA-Module aller 256 Cluster eines Chip dargestellt. Es zeigt sich, dass innerhalb eines Intervalls von ca. 25 bis 30 mV 68% der Cluster eine Antwort generieren.

Aus den Messungen folgt, dass das globale Einstellen der Steuer-Spannungen unter einem Kompromiss der Funktionalität aller Cluster zu erfolgen hat. Die Spannung V_{init} wird so eingestellt, dass auch mit einem geringen Photostrom eine ausreichende Erhöhung der Potentiale der Knoten Ψ erfolgen kann, um ein Ausgangssignal zu erzeugen. Manche Flow-Transistoren der WTA-Zellen leiten bei dieser Spannung jedoch schon einen nicht zu vernachlässigenden Strom, so dass fälschlicherweise ein Ausgangssignal erzeugt wird. Dementsprechend wird die Spannung reduziert, wobei unter Auftreten eines zu geringen Photostromes oder einer zu geringen Photostrom-Differenz eventuell kein Ausgangssignal erzeugt wird, da das Potential des Knotens Ψ innerhalb der Integrationszeit nicht das notwendige Potential erreichen kann.

Die Konsequenz des globalen Mismatching der Steuerspannungen entspricht letztlich einer Verringerung der Diskriminierungs-Genauigkeit der WTA-Zellen. Die Spezifikation der Zeitdauer von 1 ms der Photostromintegration erfordert eine Spannung V_{init} , bei der unter den definierten Lichtverhältnissen in jedem Fall Ausgangssignale erzeugt werden können. Selbst bei geringen Photoströmen müssen die Potentiale der Knoten Ψ die Höhe erreichen, in der eine Unterscheidung der Ausgangsströme erfolgt. In manchen Zellen resultiert aufgrund des Mismatching eine geringere Zeit zur Diskriminierung der Photoströme, da die Entscheidungsschwelle früher erreicht wird.

Die Genauigkeit der einzelnen WTA-Module in der Detektion der Fokalkpunkt-Position unter verschiedenen Fokalkpunkt-Größen und -Intensitäten wird in den folgenden Abschnitten bewertet.

6.3 Messung der Homogenität eines Cluster

Die Genauigkeit in der Detektion der Fokalkpunkt-Position durch ein WTA-Modul wurde mithilfe eines einzelnen wohldefinierten Spot gemessen. Die in den vorangehenden Abschnitten angesprochenen Fehlüberereinstimmungen der einzelnen WTA-Zellen untereinander äußert sich in einem Fehler der Positionsdetektion, der von der Breite des Fokalkpunktes, mit der die Photodetektor-Streifen beleuchtet werden, und dem Mismatching der Zellen abhängig ist. Um diesen Fehler zu vermessen, wurde ein Spot mit einem Durchmesser von ca. $15 \mu\text{m}$ erzeugt, der mit einer konstanten Geschwindigkeit über die 19 Photodetektor-Streifen einer Detektionsrichtung des Cluster bewegt wurde. Unter idealen Eigenschaften der Elemente würde jede der Zellen das Ausgangssignal für eine gleiche zeitliche Dauer aktivieren. Abbildung 6.9 stellt den verwendeten Versuchsaufbau dar.

Die Laserdiode erzeugt einen monochromatischen Lichtstrahl mit einer Wellenlänge von $\lambda = 675 \text{ nm}$. Durch den Neutralfilter lassen sich 8 Dekaden der Lichtintensität in Stufen von einer Dekade einstellen. Um die Strahlform des durch die Laserdiode erzeugten Lichtstrahls in ein annähernd gaußförmiges Profil zu überführen, wird ein Raumfilter mit einer Lochblende von $5 \mu\text{m}$ angewendet. Der austretende Strahl wird anschließend durch eine Fokussierlinse auf

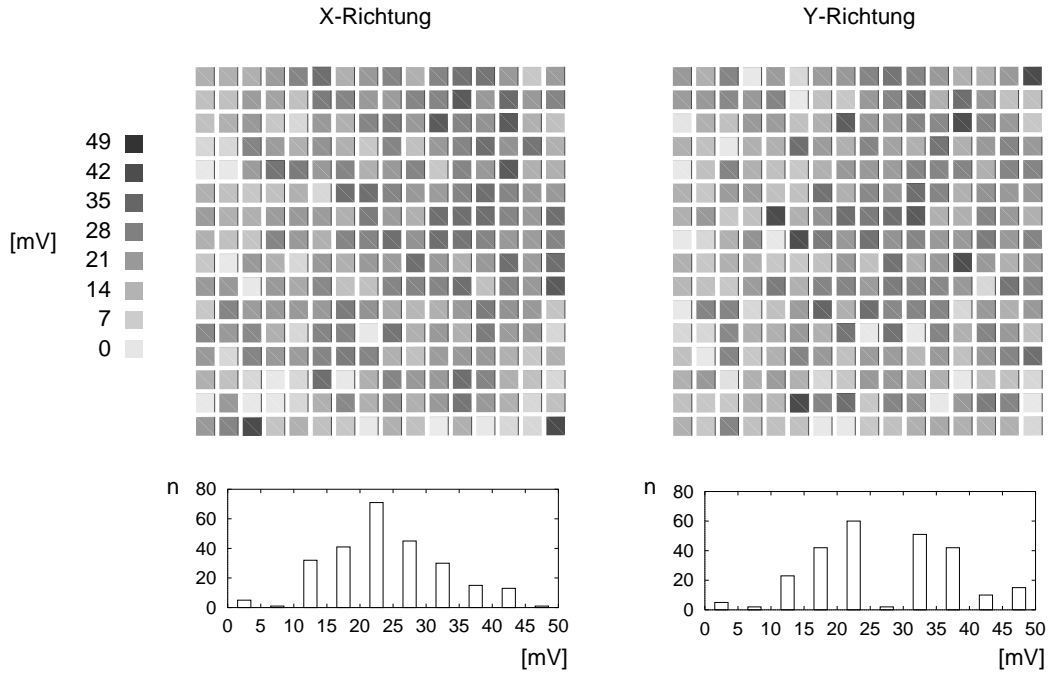


Abbildung 6.7: Verteilung der Bias-Offsets über der Cluster-Matrix des Chip I7

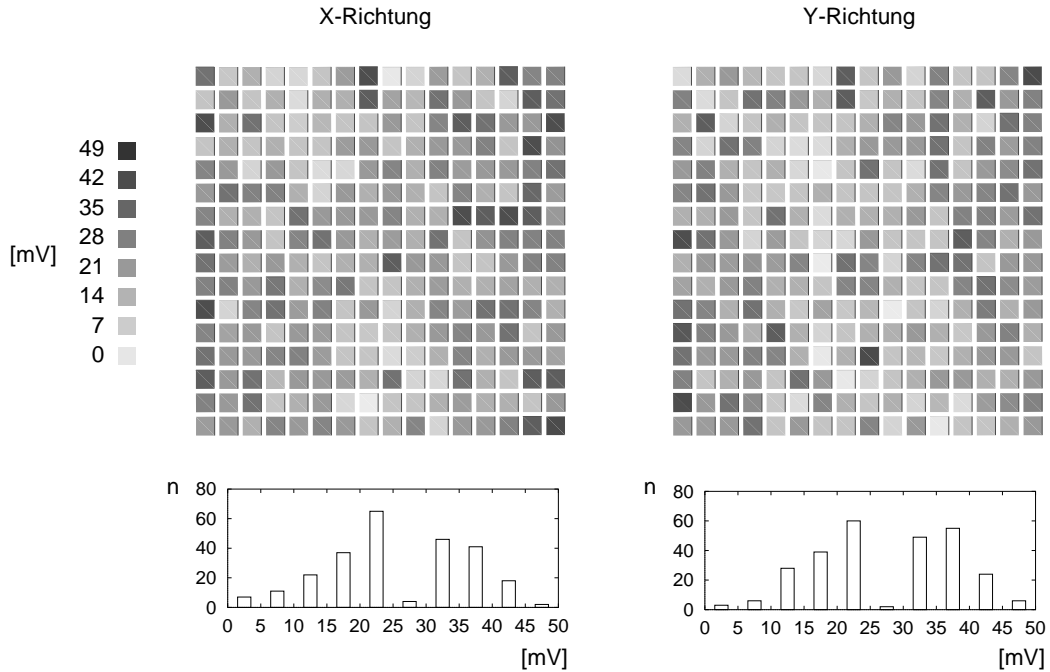


Abbildung 6.8: Verteilung der Bias-Offsets über der Cluster-Matrix des Chip I8

die Photodetektormatrix eines Cluster des ASIC fokussiert. Der Abstand der Linse von der Oberfläche des ASIC wird derart gewählt, dass die Halbwertsbreite der Intensitätsverteilung des Fokuspunktes maximal $17 \mu\text{m}$ beträgt. Unter entsprechender Positionierung des ASIC wird genau ein Photodetektorstreifen mit mindestens 68% Intensität beleuchtet. Der ASIC kann mithilfe eines rechnergesteuerten Fahrtisches mit einer absoluten Genauigkeit von $1 \mu\text{m}$ in einer zur optischen Achse des Systems vertikalen Richtung verfahren werden.

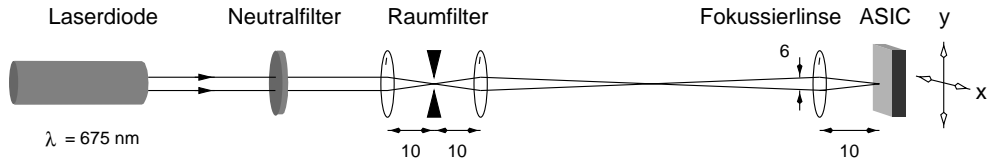


Abbildung 6.9: Der Aufbau des Messsystems zur Bestimmung der Homogenität eines Cluster

Die Lichtleistung des Spot wurde für die Messung auf den Wert von $P_{FP} \approx 2 \text{ nW}$ eingestellt. Die Integrationsdauer zur Detektion des Fokuspunktes wurde zu $t_{int} = 2 \text{ ms}$ gewählt, die Wiederholrate der Messungen betrug 250 Hz. Die Größe der Bias-Spannungen entsprachen denen von Tabelle 6.3, V_{bulk} wurde zum Abschalten der Feedback-Charakteristik auf 7 V erhöht. Anschließend wurde der ASIC über die volle Breite eines Cluster in X- und Y-Richtung mit einer Geschwindigkeit von $v = 100 \mu\text{m/s}$ gefahren und die von dem entsprechenden WTA-Modul erzeugten Bitstrings ausgelesen. Unter Verwendung der Breite eines Photodetektor-Streifens von $b_p = 17.6 \mu\text{m}$ und der Wiederholrate von 250 Hz ergibt sich bei der gewählten Geschwindigkeit des Verschiebens eine Anzahl von $E = 44$ Abtastungen, in denen der gleiche Wert der Positionsdetektion erwartet wird. Im Übergang des Fokuspunktes zwischen zwei benachbarten Photodetektor-Streifen ergibt sich ein nahezu identisches Intensitätsverhältnis, so dass die WTA-Zelle, die aufgrund des Mismatching früher einen Stromfluss des Flow-Transistors etablieren kann, "gewinnt". Dies kann geschehen, obwohl ein größerer nominaler Anteil des Photostromes in die "verlierende" Zelle fließt. Abbildung 6.10 a) gibt die von einem exemplarisch ausgewählten Cluster detektierte Position über der fortlaufenden Nummer der Abtastung wieder.

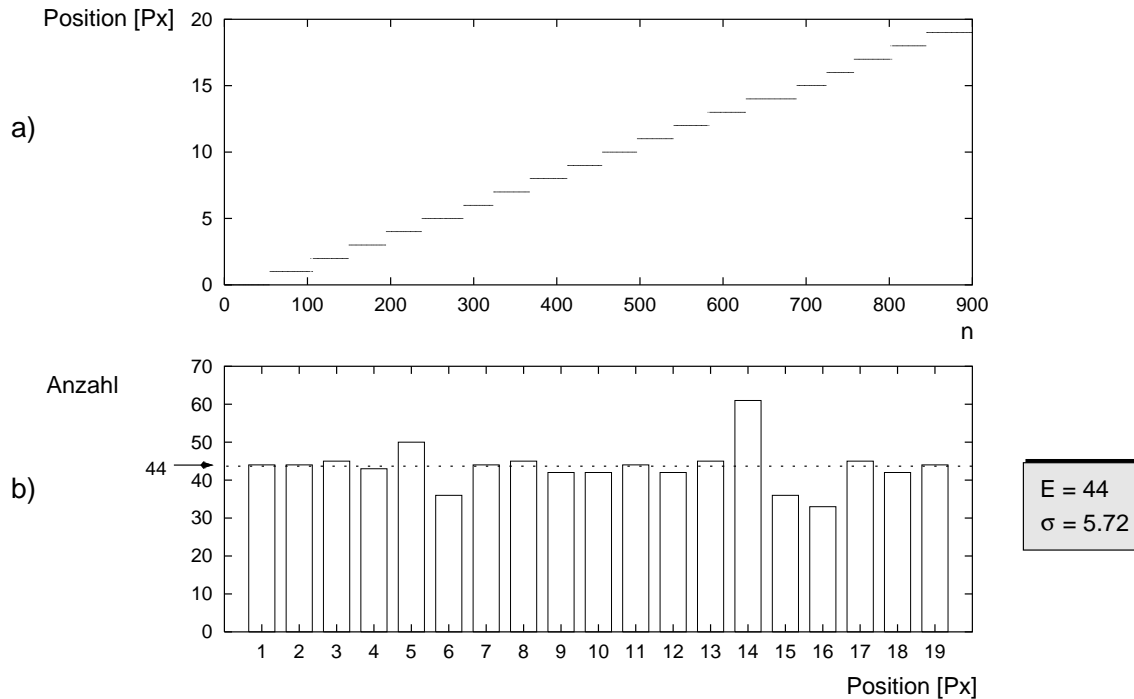


Abbildung 6.10: Ausgangssignal eines WTA-Moduls bei Verfahren eines Fokuspunktes mit einer Lichtleistung von $P_{FP} \approx 2 \text{ nW}$ und einer Größe von $D_{FP} < 17 \mu\text{m}$ über alle Zellen des Moduls. In dem unteren Diagramm ist die Anzahl der detektierten Positionswerte als Histogramm aufgetragen.

In Teilbild b) ist die Verteilung der Abtastwerte auf die einzelnen detektierten Positionen gezeigt. Für die beiden äußeren Positionen 1 und 19 wurde der Erwartungswert von 44 angenommen, da aufgrund des fehlenden äußeren angrenzenden Photodetektorstreifens ein Mismatching nicht zu quantifizieren ist. Die Werte der 17 eingefassten Positionen pendeln um den Erwartungswert von 44 und zeigen durch ihre unterschiedliche Höhe das Maß des Mismatching gegenüber den beiden benachbarten Positionen an. Dieses Mismatching ist eine Konsequenz der in Abschnitt 6.1 besprochenen Fehlübereinstimmung aller identischer Transistor-Strukturen der WTA-Zellen eines WTA-Moduls und reduziert die Genauigkeit der Positionsdetektion.

In dem exemplarischen Cluster wurde für das WTA-Modul der X-Positionen eine Standardabweichung von $\sigma = 5.72$ gemessen, so dass sich ein auf den Erwartungswert normiertes prozentuales Mismatching von 13.0% ergibt. In Tabelle 6.4 werden die von jeweils 32 Clustern der X- und Y-Richtung gemessenen gemittelten Standardabweichungen zusammengefasst.

	E	σ	%
X-Richtung	44	6.99	15.9
Y-Richtung	44	5.88	13.4

Tabelle 6.4: Arithmetisches Mittel der Standardabweichungen der Fokalkpunktposition von jeweils 32 Clustern in X- und Y-Richtung

Um den Einfluss der Rückkopplung auf das Mismatching der Zellen zu untersuchen, wurde der Versuch unter Wahl einer um den Faktor 10 geringeren Lichtleistung des Spots von $P_{FP} \approx 200 \text{ pW}$ wiederholt. Der generierte Photostrom war nicht ausreichend, um Antwortsignale der WTA-Zellen ohne Feedback-Unterstützung zu ermöglichen. Die Spannung V_{bulk} wurde daher auf 6.1 V eingestellt, um die Detektion der Fokalkpunkt-Position durch Unterstützung durch die Feedback-Ströme zu erlauben. Das Ergebniss der Messung ist in Abbildung 6.11 zu sehen.

Teilbild a) zeigt die von dem WTA-Modul des Cluster detektierte Position über der fortlaufenden Nummer der Abtastung, Teilbild b) die ermittelte Verteilung der detektierten Positionen. Es zeigt sich, dass durch die Benutzung der Feedback-Ströme eine Erhöhung des Mismatching stattfindet. Die Standardabweichung in der Anzahl der detektierten Positionen erhöht sich von $\sigma = 5.72$ auf $\sigma = 13.3$. Das Mismatching der Zelle im Feedback-Betrieb beträgt daher ca. 30%. Dieser Wert zeigt den durch den Feedback-Betrieb zusätzlich entstehenden Anteil von Mismatching in Höhe von 17%. In Tabelle 6.5 werden die von jeweils 32 Clustern der X- und Y-Richtung gemessenen gemittelten Standardabweichungen zusammengefasst.

	E	σ	%
X-Richtung	44	13.89	31.6
Y-Richtung	44	12.34	28.0

Tabelle 6.5: Arithmetisches Mittel der Standardabweichungen der Fokalkpunktposition von jeweils 32 Clustern in X- und Y-Richtung im Feedback-Modus

Die Messungen des Mismatching erfolgten unter Verwendung eines definierten Fokalkpunktes. Die Halbwertsbreite der Intensitätsverteilung war kleiner als $17 \mu\text{m}$, so dass maximal 2 Photodetektorstreifen gleichzeitig beleuchtet werden konnten. Bei Verwendung des Linsenarray werden unter Annahme von Beugungsbegrenzung Fokalkpunkt-Durchmesser von mindestens $126 \mu\text{m}$ erzeugt, die nach Messungen (siehe Abbildung 1.11) auch bis zu $200 \mu\text{m}$ betragen können. Nach

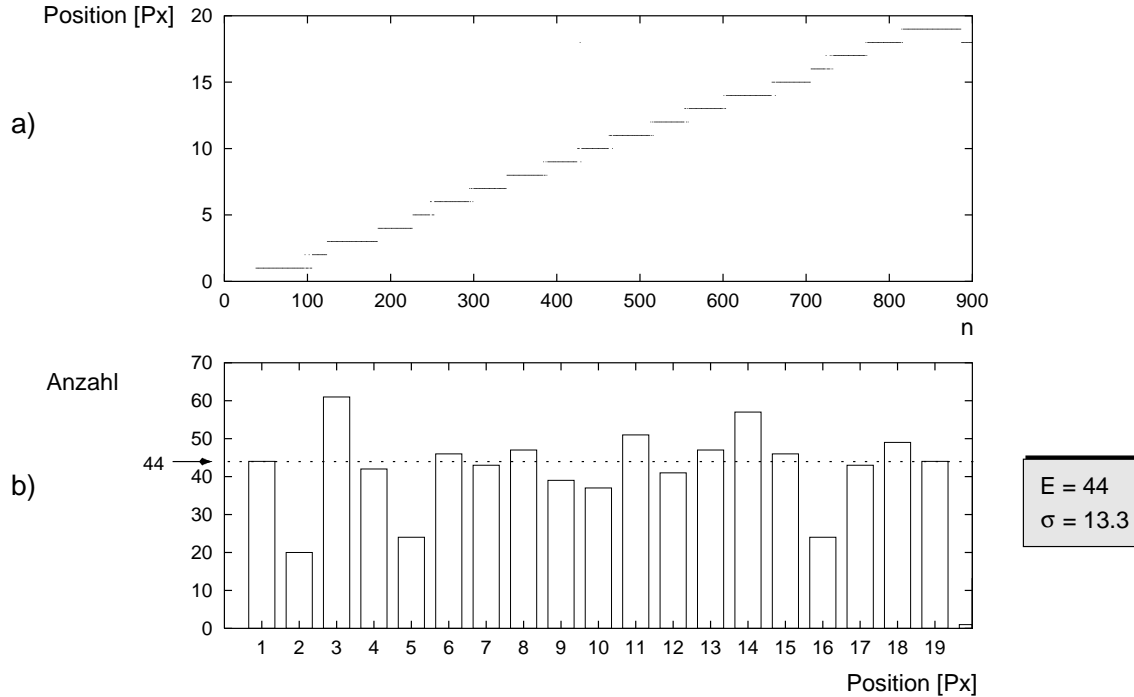


Abbildung 6.11: Ausgangssignal eines WTA-Moduls bei Verfahren eines Fokalpunktes mit einer Lichtleistung von $P_{FP} \approx 200 \text{ pW}$ und einer Größe von $D_{FP} < 17 \mu\text{m}$ über alle Zellen des Moduls unter Aktivierung der positiven Rückkopplung.

den Abschätzungen von Abschnitt 1.6.2 ist bei diesem Fokalkpunkt-Durchmesser eine Diskriminierungsgenauigkeit von 8% erforderlich, um die Position des Fokalpunktes mit der Genauigkeit einer Pixelbreite zu bestimmen. Dieser Wert unterschreitet das gemessene Mismatching der Zellen von 13 bis 15%, so dass bei Verwendung der Fokalpunkte des Linsenarray aufgrund der durch den Prozess entstehenden Unterschiedlichkeiten der WTA-Zellen ein Fehler in der Positionsdetektion entsteht.

Um Aussagen über die Diskriminierungsgenauigkeit der Positionsdetektion unter Verwendung der Linsen des Array machen zu können, wurde der Versuchsaufbau von Abbildung 6.12 verwendet.

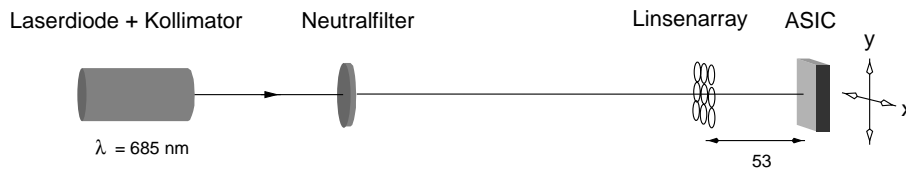


Abbildung 6.12: Der Aufbau des Messsystems

Der Kollimator bündelt den Lichtstrahl einer Laserdiode mit der Wellenlänge von $\lambda = 685 \text{ nm}$. Die Intensitätsverteilung des Strahls ist annähernd gaußförmig mit einer Halbwertsbreite von $\sigma = 15 \text{ mm}$. Die parallele Wellenfront des Kollimators, die durch den Raumfilter in ihrer Intensität abgeschwächt werden kann, durchdringt das Linsenarray und erzeugt die charakteristische Fokalkpunkt-Matrix eines Hartmann-Shack-Sensors. Der ASIC ist im Zentrum der optischen

Achse des Systems angeordnet und detektiert die Positionen der Fokalfunkte von 16×16 Linsen des Array. Die Intensität der Fokalfunkte der äußeren Cluster der Cluster-Matrix des ASIC gegenüber den zentralen Fokalfunkten unterscheidet sich aufgrund der Fläche von $6.4 \times 6.4 \text{ mm}^2$ des ASIC und dem Gaußprofil des Laserstrahls mit $\sigma = 15 \text{ mm}$ um ca. 9%. Dieser Unterschied wird vernachlässigt. Die vor dem ASIC einfallende Lichtintensität beträgt ca. 0.1 W/m^2 und erzeugt eine Lichtleistung der Fokalfunkte von $P_{FP} \approx 10 \text{ nW}$.

Es wurden die Einstellungen der Bias-Spannungen von Tabelle 6.3 gewählt, V_{bulk} wurde für die Messung zuerst auf 7 V erhöht, um das Feedback der Zellen abzuschalten. Der ASIC wurde in die X- und Y-Richtung über die Breite eines Cluster verfahren. In Abbildung 6.13 a) ist die von dem WTA-Modul des exemplarischen Cluster detektierte Position über der fortlaufenden Nummer der Abtastung, in Teilbild b) die ermittelte Verteilung der detektierten Positionen dargestellt.

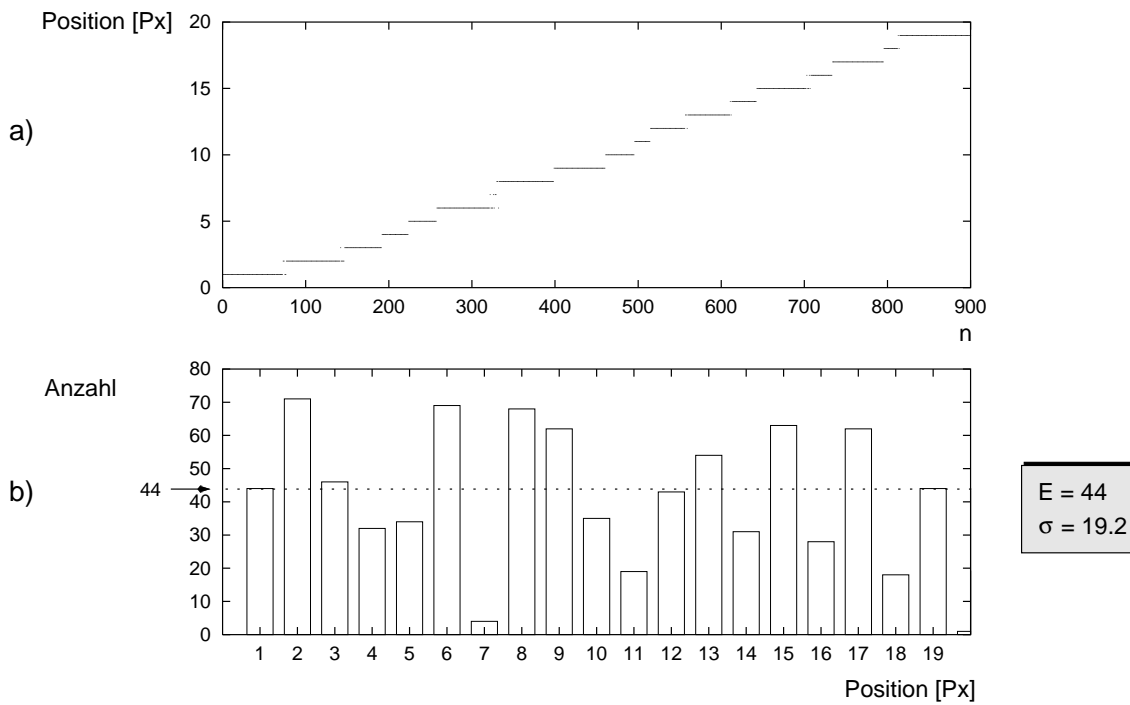


Abbildung 6.13: Ausgangssignal eines WTA-Moduls bei Verfahren eines Fokalfunktes mit einer Lichtleistung von $P_{FP} \approx 10 \text{ nW}$ und einer Größe von $D_{FP} \approx 180 \mu\text{m}$ über alle Zellen des Moduls.

Der Verlauf der Messkurve von Teilbild a) zeigt die Konsequenz des Mismatching der WTA-Zellen bei Beleuchtung der Photodetektor-Streifen mit einem mehrere Streifen umfassenden Fokalfunkt-Durchmesser. Es kann vorkommen, dass einzelne Zellen aufgrund der erhöhten Sensitivität der benachbarten Zellen und des geringen Unterschiedes der generierten Photoströme seltener ansprechen. Ein solcher Fall ist in dem exemplarischen Cluster an der Position 7 vorhanden. Die Standardabweichung in der Anzahl der detektierten Positionen des WTA-Moduls beträgt $\sigma = 19.2$, das Mismatching ca. 44%. In Tabelle 6.6 werden die von allen 256 Clustern des *HSSX* der X- und Y-Richtung gemessenen gemittelten Standardabweichungen zusammengefasst.

Die Messungen wurden für eine verringerte Lichtleistung der Fokalfunkte von $P_{FP} \approx 1 \text{ nW}$ wiederholt. Um die Detektion der Fokalfunkte zu ermöglichen, wurde der Feedback-Modus

	E	σ	%
X-Richtung	44	26.1	59.3
Y-Richtung	44	29.8	67.7

Tabelle 6.6: Arithmetisches Mittel der Standardabweichungen der Fokalfunktposition von jeweils 256 Clustern in X- und Y-Richtung unter Verwendung von Fokalfunkten mit einem Durchmesser von $D_{FP} \approx 180 \mu m$, der 8 bis 10 Photodetektor-Streifen bedeckt

der WTA-Module durch Absenken der Spannung V_{bulk} auf 5.85 V aktiviert. Das Ergebniss der Messung ist für das exemplarische Cluster in Abbildung 6.14 dargestellt.

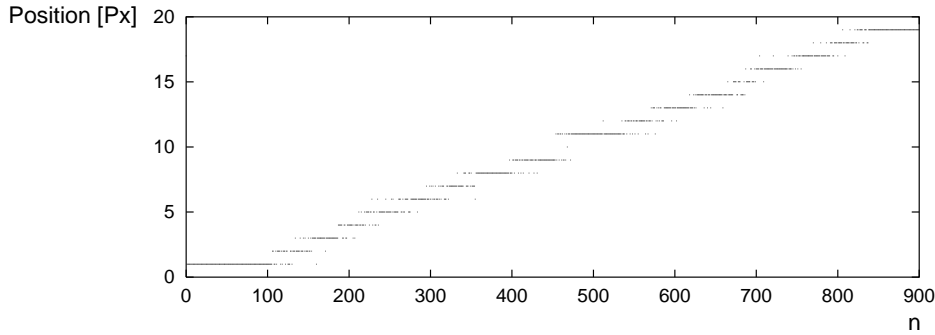


Abbildung 6.14: Ausgangssignal eines WTA-Moduls bei Verfahren eines Fokalfunktes mit einer Lichtleistung von $P_{FP} \approx 1 nW$ und einer Größe von $D_{FP} \approx 180 \mu m$ über alle Zellen des Moduls unter Aktivierung der positiven Rückkopplung

Bei der Fokalfunktintensität von $J_{FP} = 1 \frac{nW}{Spot}$ und dem Fokalfunkt-Durchmesser von $D_{FP} \approx 200 \mu m$ resultiert ein Rauschen der Positionsdetektion, das sich in einem Wechsel der für eine Position des Fokalfunktes detektierten Positionswerte um bis zu 2 Pixel äußert. Dieser Sachverhalt ist in der Abbildung 6.14 in dem scheinbaren Übereinanderliegen von Positionswerten zu erkennen.

Die Ursache des Positionsdetektions-Unsicherheit ist in der Statistik der Photostromgeneration zu finden. Nach den Ausführungen von Abschnitt *Auswirkungen von Mismatching und Rauschen* des Kapitel 3 wird durch das Schrotrauschen der Photodetektoren eine statistische Verteilung der Spannungsdifferenzen der Knoten Ψ der beiden beteiligten WTA-Zellen verursacht. Bei den aufgrund der schwachen Lichtleistung der Fokalfunkte geringen Photoströmen und der entsprechend geringen Differenz der Photoströme wird diejenige der beiden WTA-Zellen "gewinnen", die während der Integrationsphase insgesamt mehr Ladungsträger in der parasitären Kapazität des Knotens Ψ akkumulieren konnte. Durch die Feedback-Ströme wird die Differenz der durch das Schrotrauschen erzeugten Spannungsdifferenzen der Knoten Ψ verstärkt und es entsteht eine Unsicherheit in der Bestimmung der Fokalfunkt-Position. Diese Unsicherheit wurde bei verschiedenen Integrationszeiten für die beiden Chips I7 und I8 quantifiziert.

6.4 Quantifizierung des Positionsrauschens

Zur Messung der Unsicherheit in der Positionsbestimmung der Fokalfunkte wurde mit dem Messsystem aus Abbildung 6.12 eine parallel zur optischen Achse stehende Referenzwellen-

front mit einer zu vernachlässigenden Restwelligkeit der Phase von verschiedenen Chips des Wellenfrontsensor-ASIC vermessen. Dabei wurden die einzelnen Biasspannungen zur Steuerung der Komponenten der WTA-Module³ des ASIC derart eingestellt, dass jedes Cluster ein Detektionssignal erzeugt. Durch die Wahl der Spannung V_{init} wurde die Startspannung der Potentiale der Knoten Ψ_i auf 0.7 V eingestellt. Die Spannung V_{src} wurde auf 0.82 V eingestellt, was nach den Ergebnissen der Parameter-Messungen der Chips zu Beginn des Kapitels einem Source-Strom der WTA-Zellen von ca. 60 nA entspricht (vergleiche Abbildung 6.3 $W/L=10/6$). Die Spannungen V_{bout} und V_{cas} betrugen 0.75 V und 0.74 V. Der Strom der Dummy-Zellen wurde durch Wahl von $V_{\text{dmy}} = 5 \text{ V}$ derart reduziert, dass die durch die entsprechenden Transistoren der Dummy-Zellen erzeugten Ströme keinen Einfluss auf die restlichen Zellen der WTA-Module haben konnten. Der Feedback-Strom der WTA-Zellen wurde durch Einstellen der Spannung V_{blk} auf 7 V ebenfalls abgeschaltet.

In dieser Einstellung wurden die Photodetektor-Matrizen der Cluster mit den Fokalfunkten beleuchtet. Die WTA-Zelle n , die den höchsten Eingangsstrom des Photodetektor-Streifens erfährt, generiert ein gegenüber den restlichen Zellen des Moduls erhöhtes Potential des Knotens Ψ_n . Der Flow-Transistor M_{Fn} übernimmt den größten Anteil des Stromes I_{src} , der über den Stromspiegel in den Ausgangsbuffer gespiegelt wird. Bei ausreichender Höhe des Ausgangs-Stromes des Stromspiegels ergibt sich eine logische 1 am Ausgang der WTA-Zelle, die nach Aktivieren der Spannung V_{set} in die Schieberegisterzelle übernommen wird. Bei zu großer Wahl des Stromes I_{src} durch die Spannung V_{src} und zu geringer Wahl der Spannung V_{bout} kann es vorkommen, dass zwei oder mehrere Zellen eines WTA-Moduls eine logische 1 erzeugt. Dies wird durch den Kompressor mit dem Double- oder dem Error-Bit angezeigt. Sollte der Strom aller Flow-Transistoren eines WTA-Moduls zu gering sein, wird von keiner Zelle eine logische 1 übergeben, was durch das Dark-Bit des Kompressors angezeigt wird.

Bei einer Integrationsdauer von $t_{\text{int}} = 1 \text{ ms}$ reichte die Lichtleistung der Fokalfunkte von $J_{\text{FP}} \approx 1 \frac{\text{nW}}{\text{Spot}}$ nicht aus, um durch eine WTA-Zelle ohne Verwendung des Feedback-Stromes ein Ausgangssignal zu generieren. Daher wurde die Spannung V_{bulk} reduziert, um eine Unterstützung der Photoströme durch die Feedback-Ströme zu etablieren. Bei einer Spannung von $V_{\text{bulk}} = 5.85 \text{ V}$ konnte ein ausreichender Feedback-Strom erzeugt werden, um durch die Zellen Ausgangssignale zu erzeugen. Durch Einstellen der Spannungen V_{init} , V_{bout} und V_{bulk} wurde ein Arbeitspunkt des ASIC definiert, zu dem alle Cluster ein Positions-Signal generierten, also jeweils das Dark-Bit gelöscht war, und die gelöschten Double- und Error-Bits die Eindeutigkeit der detektierten Positionen anzeigten.

Aufgrund der breiten Gestalt der Fokalfunkt-Intensitätsverteilung und der daraus resultierenden geringen Differenz der Photoströme und der Unsicherheit der Photoströme aufgrund des Schrotrauschens waren die Positionssignale zeitlich nicht konstant, sondern umfassten in zeitlicher Abfolge einige aufeinanderfolgende Werte. Dies stellte sich als Unsicherheit der Positions-Detektion dar und wird im Folgenden als *Positionsrauschen* bezeichnet.

Durch Wahl einer längeren Integrationsdauer von 2 oder 4 ms konnte bei gleichzeitiger Erhöhung der Spannung V_{bulk} zur Reduzierung der Feedback-Ströme ein geringeres Positionsrauschen gemessen werden. In den Abbildungen 6.15 und 6.16 ist das Positionsrauschen der Chips I7 und I8 unter Wahl der drei Integrationszeiten von $t_{\text{int}} \in \{1 \text{ ms}, 2 \text{ ms}, 4 \text{ ms}\}$ unterteilt in X- und Y-Richtung dargestellt. Die Graphen stellen Histogramme dar, deren Höhenwerte die Differenz zwischen maximaler und minimaler detektierter Position des entsprechenden Cluster von 100 Messungen in Einheiten von Pixel über den 16×16 Clustern des ASIC angeben. Unter den Graphen ist der Durchschnitt des Positionsrauschens aller 256 Cluster angegeben.

³Siehe Abbildung 3.24

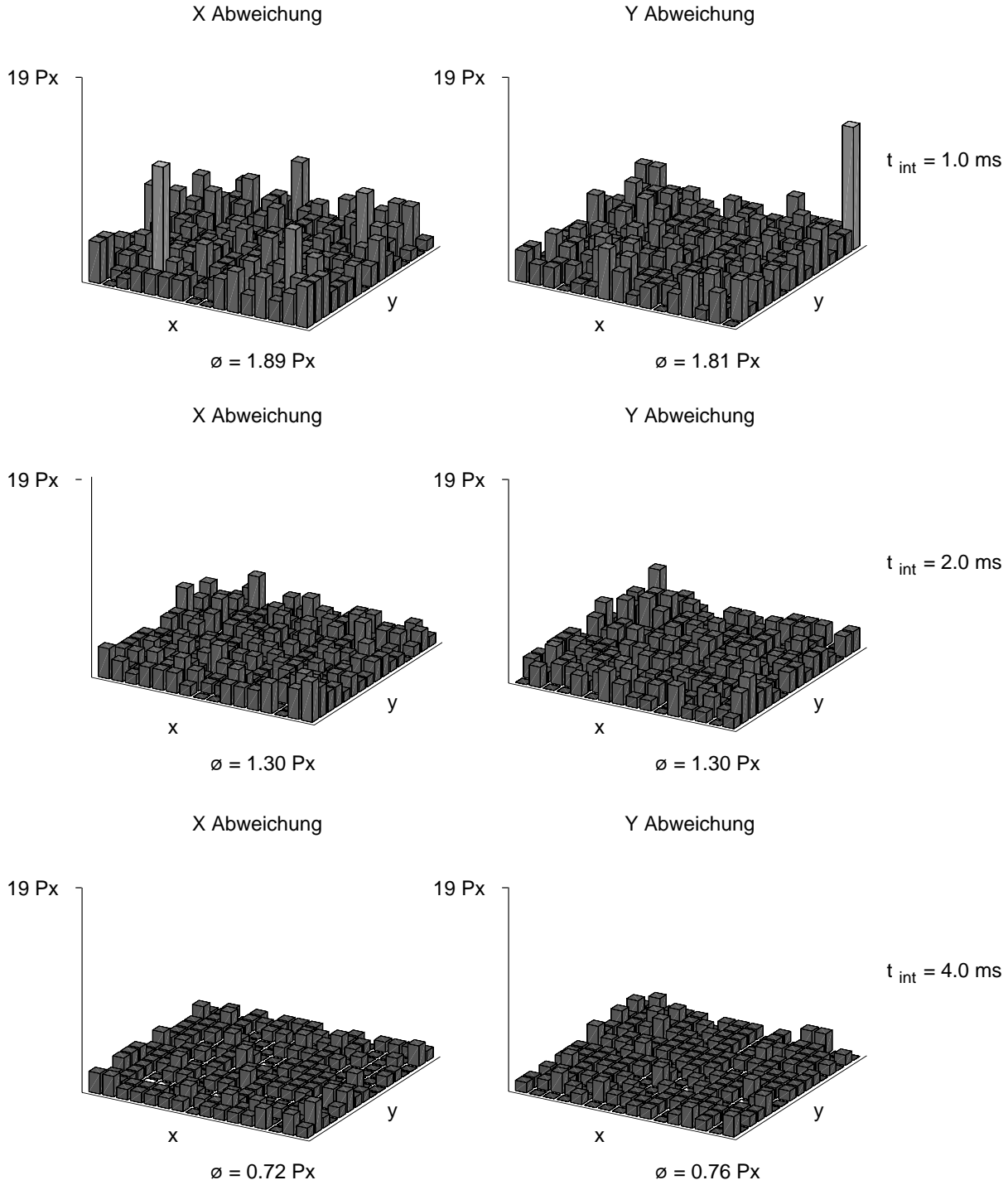


Abbildung 6.15: Rauschen in der Positionsbestimmung der Fokalkpunkt-Position unter verschiedenen Integrationsdauern des Chip I7

Es wurde festgestellt, dass die Größe des Positionsrauschens von der Intensität des Lichtes, der zeitlichen Dauer der Detektion und der Einstellung der Spannung V_{bulk} abhängig ist. Unter Wahl einer Integrationszeitdauer von $t_{\text{int}} = 1 \text{ ms}$ bei $P_{\text{FP}} \approx 1 \text{ nW}$ ergab sich für Chip I7 ein in X- und Y-Richtung durchschnittliches Rauschen der Positionsdetektion von ca. 2.13 Pixel. Die

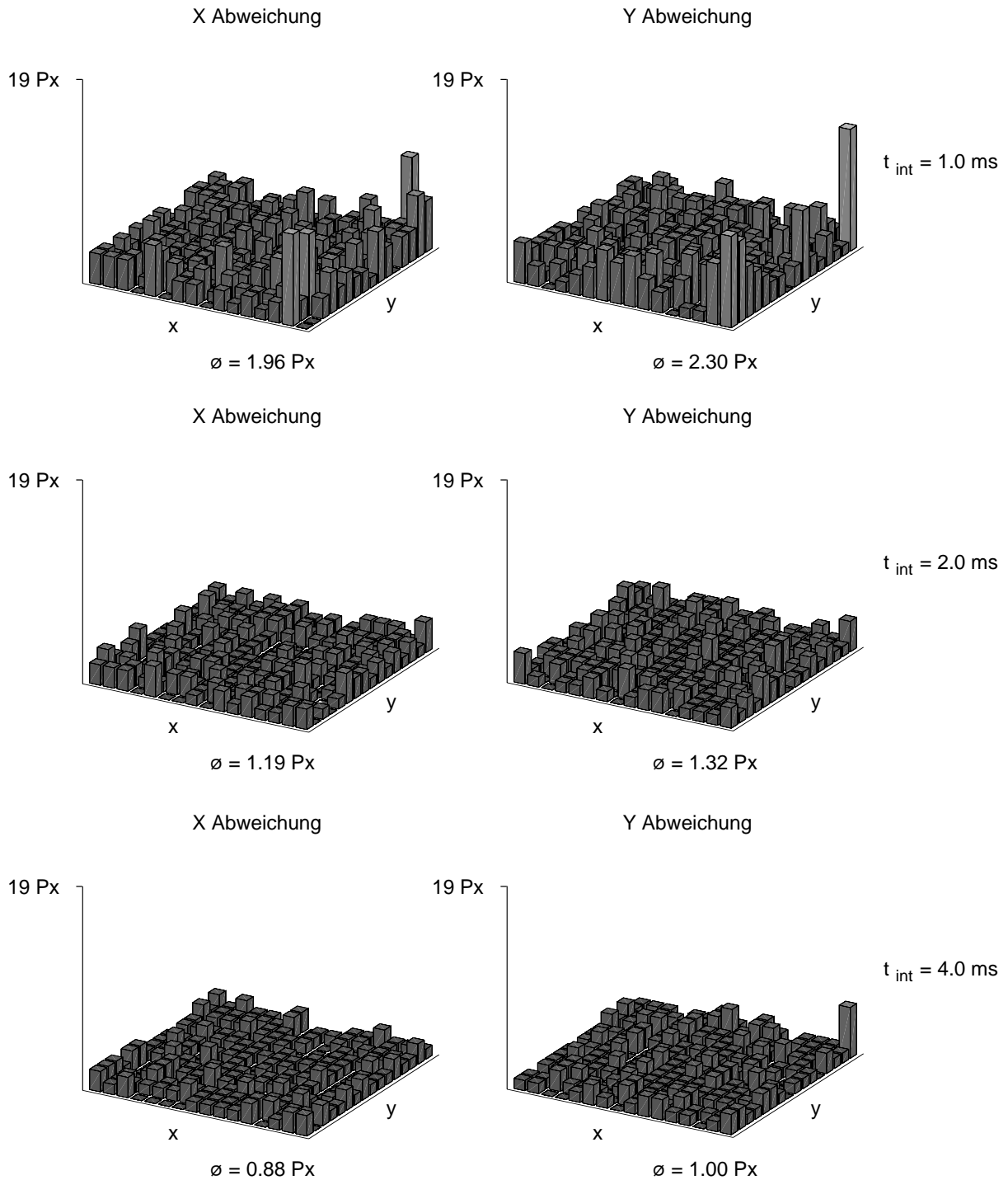


Abbildung 6.16: Rauschen in der Fokuspunkt-Positionsbestimmung der 16×16 Cluster bei verschiedenen Integrationsdauern des Chip I8

Messungen von Chip I8 ergab Werte, die ein Rauschen der Positionsdetektion von ca. 1.85 Pixel zeigen. Unter Wahl einer längeren Dauer der Integration und erhöhter Spannung von V_{bulk} wurde das Positionsruschen von I7 und I8 bei gleicher Lichtleistung mit $t_{int} = 2$ ms auf 1.3 und 1.25 Pixel, mit $t_{int} = 1$ ms auf 0.74 und 0.94 Pixel reduziert.

6.5 Messung der optischen Aberrationen von Test-Linsen

Um die Qualität der Positionsdetektion in Bezug auf die Rekonstruktion der Wellenfront durch Berechnung der Zernike-Koeffizienten abzuschätzen, wurden mehrere optische System auf die Quantität von optischen Aberrationen untersucht. Die optischen Medien von 4 verschiedenen Sehhilfen wurden verwendet, da der dynamische Bereich der optischen Aberrationen im Einklang mit den in Kapitel 1.6 benannten Spezifikationen der Fokalkpunkt-Positionsdetektion liegt. Das Maß der optischen Aberration war für die vermessenen Objekte in Form von Spezifikationen der optischen Aberrationen definiert und konnte als Kontrolle der erzielten Messungen verwendet werden.

Das Messsystem aus Abbildung 6.12 wurde gemäß Abbildung 6.17 erweitert. Zwischen dem Neutralfilter und dem Linsenarray wurden die Linsen der verschiedenen Sehhilfen angeordnet und die durch das Linsenarray erzeugten Fokalkpunkt-Verschiebungen detektiert. Die Daten von jeweils 100 Messungen wurden in Echtzeit von einem PC durch Least-Square-Fit in Zernike-Koeffizienten umgerechnet⁴. Die 4 verschiedenen Linsen werden im Folgenden als Testobjekt I-IV bezeichnet.

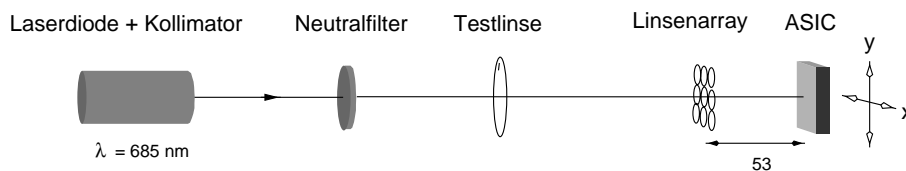


Abbildung 6.17: Der Aufbau des Messsystems zur Vermessung von Linsen mit definierten optischen Eigenschaften.

In Tabelle 6.7 sind die Ergebnisse des Least-Square-Fit von Testobjekt I unter Verwendung der Chips I7 und I8 bei drei verschiedenen Integrationszeiten von t_{int} zusammengefasst. In den Zeilen sind jeweils die durchschnittlichen 14 Zernike-Koeffizienten und darunter die dazugehörige Standardabweichung von 100 Messungen des Koeffizienten eingetragen. Die Größe der Standardabweichung reduziert sich unter Wahl einer längeren Integrationsdauer, da bei einer Verlängerung der Integrationszeit die Steuerspannung der Feedback-Ströme erhöht und das Positionsrauschen verringert werden kann.

Es zeigt sich, dass die von den beiden Chips vermessenen optischen Aberrationen annähernd die gleichen Werte einnehmen. Hervorzuheben ist der etwas größere Unterschied in der Bestimmung des Koeffizienten C_3 . Dieser definiert den Astigmatismus in 45° -Richtung und wird bei einer Fehljustage der ASICs in Bezug auf deren Rotation um die optischen Achse des Systems überhöht bewertet. Da die Chips unter gleichen Bedingungen vermessen werden sollten, wurde der Messaufbau nicht verändert. Aufgrund der leichten unterschiedlichen Verdrehung der Chips auf dem jeweiligen Chipcarrier⁵ sprach der Koeffizient C_3 entsprechend unterschiedlich an. Für Einzelmessungen wird das System derart eingerichtet, dass die Ausrichtung des Linsenarray mit der Ausrichtung der Clustermatrix des HSSX übereinstimmt.

Im weiteren Verlauf der Untersuchungen wurden durch Chip I8 die optischen Aberrationen der Testobjekte II, III und IV vermessen. Die aus den Fokalkpunkte-Positionen errechneten durchschnittlichen Zernike-Koeffizienten und die dazugehörigen Standardabweichungen sind in Tabelle

⁴Die Zeitdauer der Berechnung durch den Least-Square-Fit erforderte auf einem Pentium II-Prozessor mit 300 MHz Taktfrequenz eine Zeitdauer von ca. $280 \mu\text{s}$, gemessen unter Maskierung der Interrupts und Benutzung des CPU-eigenen 1.2 MHz-Zählers

⁵Keramik-Träger der Chips zur Montage auf PCBs

Chip I7: Zernike-Koeffizienten und Standardabweichung von Testobjekt I														
$t_{int} = 1\text{ ms}$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.36	-1.26	-0.51	-0.18	-1.10	-0.04	-0.05	0.02	-0.02	-0.02	0.02	-0.02	0.02	0.02
σ_j	0.04	0.04	0.13	0.02	0.05	0.06	0.04	0.03	0.07	0.04	0.02	0.02	0.03	0.04
$t_{int} = 2\text{ ms}$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.07	-1.16	-0.48	-0.17	-1.12	0.01	-0.03	0.02	-0.05	-0.01	0.01	0.00	0.01	-0.01
σ_j	0.02	0.02	0.07	0.01	0.03	0.04	0.02	0.02	0.04	0.02	0.01	0.01	0.02	0.03
$t_{int} = 4\text{ ms}$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.05	-1.16	-0.50	-0.17	-1.14	0.03	-0.02	0.01	-0.04	0.00	0.01	0.00	0.01	0.01
σ_j	0.02	0.02	0.05	0.01	0.03	0.03	0.02	0.02	0.03	0.02	0.01	0.01	0.02	0.02
Chip I8: Zernike-Koeffizienten und Standardabweichung von Testobjekt I														
$t_{int} = 1\text{ ms}$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	-0.38	-0.82	-0.59	-0.19	-1.14	-0.07	0.03	-0.03	0.08	-0.01	-0.02	0.00	-0.04	-0.07
σ_j	0.04	0.04	0.11	0.02	0.04	0.05	0.03	0.03	0.06	0.04	0.02	0.02	0.03	0.04
$t_{int} = 2\text{ ms}$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	-0.71	-0.90	-0.62	-0.18	-1.16	-0.03	-0.02	-0.02	0.01	0.02	0.00	-0.01	0.00	0.02
σ_j	0.02	0.02	0.06	0.01	0.03	0.03	0.02	0.02	0.03	0.02	0.01	0.01	0.01	0.02
$t_{int} = 4\text{ ms}$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	-0.86	-0.94	-0.53	-0.19	-1.16	-0.04	-0.03	-0.03	0.02	0.01	0.01	0.00	0.01	0.02
σ_j	0.02	0.01	0.04	0.01	0.02	0.02	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.02

Tabelle 6.7: Durchschnitt und Standardabweichung der Zernike-Koeffizienten von 100 Wellenfrontinformationen des Testobjekt I in Einheiten von μm . Die Werte wurden von den Chips I7 und I8 gemessen. Die hervorgehobenen Werte C_3 , C_4 und C_5 drücken Astigmatismus in 90°-Richtung, Defokus und Astigmatismus in 45°-Richtung aus.

6.8 aufgelistet. Es ist anzumerken, dass die Positionsdetektion des *HSSX* aufgrund der Diskretisierung der tatsächlichen Position in X- Y-Abschnitte mit einer Breite von minimal $\Delta x_d = 17\text{ }\mu\text{m}$ innerhalb des dynamischen Bereichs der Positionsdetektion von $340\text{ }\mu\text{m}$ einen absoluten Fehler der Zernike-Koeffizienten-Berechnung durch den Least-Square-Fit von 0.06 dpt verursacht, siehe Abschnitt 1.6.1. Bei einem Positionsrauschen von 2 Pixeln, entsprechend $\Delta x_d = 34\text{ }\mu\text{m}$, steigt der absolute mögliche Fehler auf 0.12 dpt an.

Die aus den Positionswerten der Fokalfunkte errechneten Zernike-Koeffizienten stellen nach Abschnitt 1.4.1 eine mathematische Beschreibung der Wellenfront dar. Eine für die qualitative Beurteilung der optischen Aberrationen der Testobjekte bessere Darstellung erfolgt durch die Angabe von Defokus (A_S) und Astigmatismus mit Winkel (A_Z , Θ_Z)⁶ in Einheiten von Dioptrien, respektive Grad. Die Umrechnungsformeln für diese Werte aus den Zernike-Koeffizienten, die auf eine Apertur der Fokalfunkte-Matrix von 6.4 mm bezogen sind, werden durch die Gleichungen 6.1 angegeben.

⁶Diese Typen von Aberrationen werden in der Ophthalmologie als sphärischer Fehler und zylindrischer Fehler unter einem Winkel referenziert und im Allgemeinen als Beschreibung der Korrektur von Sehhilfen verwendet

Chip I8: Zernike-Koeffizienten und Standardabweichung von Testobjekt II														
$t_{int} = 1\ ms$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.01	-0.25	0.81	0.67	-2.92	-0.02	-0.06	0.00	0.01	0.06	-0.06	-0.03	0.06	0.01
σ_j	0.04	0.04	0.12	0.03	0.05	0.08	0.04	0.04	0.06	0.05	0.03	0.02	0.03	0.06
$t_{int} = 2\ ms$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.04	-0.40	0.69	0.67	-2.95	-0.02	-0.08	0.03	-0.02	0.10	-0.05	-0.04	0.09	0.01
σ_j	0.03	0.02	0.12	0.03	0.03	0.06	0.05	0.03	0.04	0.06	0.03	0.02	0.03	0.05
$t_{int} = 4\ ms$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.19	-0.24	0.61	0.69	-2.93	-0.04	-0.09	-0.01	0.02	0.11	-0.05	-0.04	0.07	0.01
σ_j	0.02	0.02	0.10	0.03	0.03	0.05	0.05	0.02	0.03	0.05	0.02	0.02	0.03	0.05
Chip I8: Zernike-Koeffizienten und Standardabweichung von Testobjekt III														
$t_{int} = 1\ ms$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.53	0.24	0.33	1.83	0.06	0.13	-0.08	-0.25	0.00	-0.28	-0.13	-0.13	-0.16	-0.15
σ_j	0.04	0.05	0.14	0.03	0.06	0.07	0.04	0.05	0.09	0.05	0.03	0.02	0.05	0.05
$t_{int} = 2\ ms$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.37	0.23	0.03	1.92	0.02	0.01	-0.02	0.01	-0.14	0.01	0.00	-0.01	-0.04	0.04
σ_j	0.03	0.02	0.07	0.01	0.03	0.03	0.02	0.02	0.04	0.02	0.01	0.01	0.02	0.03
$t_{int} = 4\ ms$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.36	0.21	0.00	1.93	0.01	0.03	0.00	0.00	-0.12	-0.01	0.00	0.00	-0.02	0.04
σ_j	0.02	0.02	0.05	0.01	0.02	0.02	0.01	0.01	0.03	0.01	0.01	0.01	0.01	0.02
Chip I8: Zernike-Koeffizienten und Standardabweichung von Testobjekt IV														
$t_{int} = 1\ ms$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.58	-0.48	0.30	-0.06	-2.59	-0.02	0.02	-0.01	0.02	-0.06	-0.01	-0.01	-0.01	0.00
σ_j	0.06	0.03	0.10	0.02	0.04	0.05	0.02	0.03	0.05	0.03	0.02	0.01	0.02	0.04
$t_{int} = 2\ ms$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.61	-0.47	0.22	-0.07	-2.64	-0.02	0.01	0.00	-0.01	-0.02	0.00	-0.01	-0.01	0.01
σ_j	0.05	0.02	0.08	0.01	0.02	0.03	0.01	0.03	0.04	0.03	0.01	0.01	0.02	0.03
$t_{int} = 4\ ms$														
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
C_j	0.58	-0.49	0.14	-0.06	-2.64	-0.02	0.00	0.01	-0.03	0.01	-0.01	-0.01	0.00	0.03
σ_j	0.06	0.02	0.07	0.01	0.02	0.02	0.01	0.02	0.03	0.02	0.01	0.01	0.02	0.02

Tabelle 6.8: Durchschnitt und Standardabweichung der Zernike-Koeffizienten von 100 Wellenfrontinformationen der Testobjekte II, III und IV in Einheiten von μm . Die Werte wurden von Chip I8 gemessen.

$$\begin{aligned}
A_S &= \frac{1}{2.56} C_4 \text{ [dpt]} \\
A_Z &= \frac{1}{2.56} \sqrt{C_3^2 + C_5^2} \text{ [dpt]} \\
\Theta_Z &= \frac{1}{2} \arctan \frac{C_3}{C_5} \text{ [}^\circ\text{]}
\end{aligned} \tag{6.1}$$

In Tabelle 6.9 sind die für die 4 Testobjekte aus den Zernike-Koeffizienten von Tabelle 6.7 und 6.8 berechneten Werte der optischen Aberrationen von Defokus und Astigmatismus eingetragen. Zusätzlich zu der Standardabweichung der Werte der 100 Messungen von Chip 8 ist der absolute Fehler der Berechnung durch die Diskretisierungsungenauigkeit dargestellt. Die Genauigkeit unter Wahl der verschiedenen Integrationszeiten wird durch das entsprechende Positionsrauschen definiert, das in Abbildung 6.16 dargestellt ist. Für die Werte $t_{int} \in \{1\text{ ms}, 2\text{ ms}, 4\text{ ms}\}$ beträgt das über die beiden Detektionsrichtungen gemittelte Positionsrauschen 2.13, 1.26 und 0.94 Pixel, zuzüglich 1 Pixel Quantisierungsfehler. Bezogen auf die 19 Pixel dynamischen Bereich, entsprechend ± 1 dpt optischer Aberration, beträgt der maximale absolute Fehler in der Bestimmung der optischen Aberrationen A_S und A_Z ± 0.16 , ± 0.12 und ± 0.1 dpt.

Chip I8: Optische Aberrationen von Testobjekt I								
	Messungen					Spezifikation		
t_{int}	A_S	A_Z	Fehler	Θ_Z [°]	Fehler	A_S	A_Z	Θ_Z [°]
1 ms	-0.074 ± 0.008	0.501 ± 0.046	± 0.16	-14 ± 3	± 5	-0.25	-0.5	-10
2 ms	-0.070 ± 0.004	0.514 ± 0.026	± 0.12	-14 ± 2	± 3			
4 ms	-0.074 ± 0.004	0.498 ± 0.017	± 0.10	-12 ± 1	± 3			
Chip I8: Optische Aberrationen von Testobjekt II								
	Messungen					Spezifikation		
t_{int}	A_S	A_Z	Fehler	Θ_Z [°]	Fehler	A_S	A_Z	Θ_Z [°]
1 ms	0.262 ± 0.012	1.184 ± 0.051	± 0.16	82 ± 1	± 2	-0.25	-1.25	80
2 ms	0.262 ± 0.012	1.183 ± 0.048	± 0.12	83 ± 1	± 2			
4 ms	0.270 ± 0.012	1.169 ± 0.041	± 0.10	84 ± 1	± 1			
Chip I8: Optische Aberrationen von Testobjekt III								
	Messungen					Spezifikation		
t_{int}	A_S	A_Z	Fehler	Θ_Z [°]	Fehler	A_S	A_Z	Θ_Z [°]
1 ms	0.715 ± 0.012	0.131 ± 0.059	± 0.16	-		-0.75	0	0
2 ms	0.750 ± 0.004	0.014 ± 0.030	± 0.12	-				
4 ms	0.754 ± 0.004	0.004 ± 0.021	± 0.10	-				
Chip I8: Optische Aberrationen von Testobjekt IV								
	Messungen					Spezifikation		
t_{int}	A_S	A_Z	Fehler	Θ_Z [°]	Fehler	A_S	A_Z	Θ_Z [°]
1 ms	-0.023 ± 0.008	1.018 ± 0.042	± 0.16	87 ± 1	± 2	-0.5	-1.0	90
2 ms	-0.027 ± 0.004	1.035 ± 0.032	± 0.12	88 ± 1	± 2			
4 ms	-0.023 ± 0.004	1.033 ± 0.028	± 0.10	88 ± 1	± 1			

Tabelle 6.9: In Dioptrien angegebene vermessene Aberrationen der Testobjekte und ihre Spezifikationen. Die Werte sind durch arithmetische Mittelung aus 100 Messungen berechnet worden. Die Standardabweichung ist direkt hinter dem jeweiligen Wert angegeben. In der Spalte *Fehler* ist der aus dem Positionsrauschen und der Quantisierung theoretisch entstehende maximale absolute Fehler der Messwerte angegeben.

Die in der Tabelle eingetragenen Werte der optischen Aberrationen von Messungen und Spezifikationen verwenden gegensätzliche Vorzeichen. Dies geschieht aufgrund der Tatsache, dass der Wellenfrontsensor-ASIC die tatsächliche Korrektur der Linsen misst, während die Werte der Spezifikationen der Linsen den Fehler des zu korrigierenden Auges angeben. Das Ergebnis einer

Addition der beiden Werte der optischen Aberration von Auge und Linse sollte 0 ergeben, um eine Korrektur der Fehlsichtigkeit zu bewirken.

Nach Vergleich der vermessenen Werte mit den Spezifikationen wird festgestellt, dass die gemessenen und die spezifizierten Werte der optischen Aberrationen für die astigmatischen und defokalen Aberrationen im Allgemeinen gut übereinstimmen. Bei den Testobjekten I und IV ergaben sich jedoch Diskrepanzen der defokalen Aberrationen. Nachdem das Messverfahren überprüft und die Messungen mehrmals wiederholt wurden, wurde angenommen, dass die spezifizierten Werte der sphärischen Korrekturen nicht mit den tatsächlich vorhandenen Werten übereinstimmen. Vergleichswerte konnten zur Stützung dieser Aussage nicht hinzugezogen werden, da die Testlinsen nur in einem begrenzten Zeitraum zur Verfügung standen und daher weitere Messungen in unabhängigen Systemen nicht möglich waren.

6.6 Leistungsaufnahme des HSSX

Die Stromaufnahme des *HSSX* wurde im Hinblick auf die frequenzabhängige Leistungsaufnahme gemessen. Im statischen Betrieb resultiert ein Strom, der sich aus der Summe der I_{src} und I_{cas} aller WTA-Module zusammensetzt. Nach Messungen addieren sich die Ströme zu ca. $400 \mu\text{A}$. Die dynamischen Schieberegister und der Kompressor werden über Reset-Transistoren in einen stromlosen Zustand gehalten und benötigen bis auf den Verbrauch von Sperr-Sättigungsströmen der pn-Übergänge der verschiedenen Drain- oder Source-Diffusionsen im Signalpfad keinen nennenswerten Strom.

Im Betrieb des *HSSX* werden während einer Integrationsdauer die Potentiale der WTA-Module durch die Photo- oder Dummy-Ströme eingestellt und Ausgangssignale generiert. Anschließend erfolgt das Übertragen der Daten in die Schieberegister und deren Kompression durch den Kompressor. Diese Funktionsgruppen benötigen während ihres Betriebes einen frequenzabhängigen Strom, der in Tabelle 6.10 für verschiedene Betriebsfrequenzen eingetragen ist. Die unterste Zeile der Tabelle stellt die auf eine Versorgungsspannung von 5 V bezogen Verlustleistung des *HSSX* dar. Die Stromaufnahme des Kompressors und der Ausgangspads bei einer Frequenz des Systems von 1 kHz beträgt ca. 50% des in Kapitel 3 berechneten Wertes. Die Berechnung der Werte erfolgte unter Anwendung der angegebenen frequenzbezogenen Leistungsaufnahme der Standardzellen aus [AMS 0.6 III], die eine Worst-Case-Abschätzung darstellen. Der Prozess ergab weitestgehend *Typical-Mean*-Parameter, siehe Abschnitt 6.1, so dass die tatsächliche Stromaufnahme geringer ausfällt.

Frequenz	Hz	0	200	250	333	500	1000
I(Cluster-Matrix)	mA	0.40	0.51	0.54	0.59	0.70	1.00
I(dynamische Schieberegister)	mA	0.00	0.08	0.10	0.13	0.20	0.41
I(Kompressor + Pads)	mA	0.00	0.05	0.06	0.08	0.12	0.23
I_{tot}	mA	0.40	0.64	0.70	0.80	1.02	1.64
$P_{\text{tot}} @ 5 \text{ V}$	mW	2.0	3.2	3.5	4.0	5.1	8.2

Tabelle 6.10: Stromaufnahme der verschiedenen Funktionsgruppen des *HSSX* und gesamte Leistungsaufnahme bei verschiedenen Frequenzen

6.7 Diskussion der Messergebnisse

Die in diesem Kapitel vorgestellten Untersuchungen der Funktionalität des *HSSX* beruhen auf der Vermessung von 32 der insgesamt 180 verschiedenen Chips. Durch die Messung der Ausgangs- und der Übertragungs-Kennlinien von verschiedenen MOS-FET-Strukturen wurden zwei Exemplare ausgewählt, bei denen eine sehr gute Übereinstimmung der Drain-Ströme von identischen, im Abstand von 70 und 520 μm benachbarten Transistoren festgestellt werden konnte.

Durch Vermessung der Positionsdetektions-Einheiten der Chips wurde das Mismatching der einzelnen WTA-Module unter Verwendung von verschiedenen Spot-Größen und -Intensitäten quantifiziert. Die Auswirkungen der positiven Rückkopplung der WTA-Module zur Ermöglichung der Detektion sehr lichtschwacher Fokalfunkte konnte evaluiert werden.

Bei Verwendung eines Fokalfunktes, dessen laterale Ausdehnung maximal der Breite eines Photodetektor-Streifens entsprach und dessen Lichtleistung ca. 2 nW betrug, konnte eine Fehl-übereinstimmung der WTA-Zellen von 13.5 bis 16% gemessen werden. Bei Verringerung der Lichtleistung um den Faktor 10 auf 200 pW wurde der Feedback-Betrieb der WTA-Zellen notwendig, um definierte Ausgangssignale zu erzeugen. Durch das Mismatching der zusätzlich aktivierten MOS-FET erhöhte sich die Fehl-übereinstimmung auf 28 bis 31.6%.

Bei Verwendung eines Fokalfunktes, dessen laterale Ausdehnung ca. 200 μm und dessen Lichtleistung ca. 10 nW betrug, vergrößerte sich die Ungenauigkeit weiterhin. Aufgrund der breiten Intensitätsverteilung entstehen in den beleuchteten Detektor-Streifen geringfügige Photostrom-Unterschiede, die nach den Berechnungen von Abschnitt 1.6.2 des Kapitels 1 im Zentrum des Fokalfunktes ca. 8% betragen. Dadurch ist eine größere Positionsänderung des Fokalfunktes erforderlich, um einen für die entsprechende WTA-Zelle messbaren Stromunterschied zu generieren. Das Mismatching ohne Verwendung des Feedback-Modus erhöhte sich auf 59 bis 67%.

Die Lichtleistung des Fokalfunktes wurde auf ca. 1 nW eingestellt, und der Feedback-Modus aktiviert. Es entstand ein Positionsrauschen, dass durch die statistische Verteilung der Rauschanteile der Photostrome verursacht wird. Bei den geringen Photostrom-Differenzen, die bei der eingestellten Lichtleistung des Fokalfunktes von 1 nW in benachbarten WTA-Zellen erzeugt werden, ist ein starker Feedback-Strom notwendig, um Ausgangssignale zu erzeugen. Durch die Feedback-Ströme wird die Differenz der durch das Schrotrauschen erzeugten Spannungsdifferenzen der Knoten Ψ jedoch verstärkt, so dass je nach Anzahl von Rauschelektronen während der Integrationsphase bei geringen Photostrom-Differenzen eine Unsicherheit der Positionsdetektion entsteht.

In Versuchen wurde das Positionsrauschen für zwei Chips bei verschiedenen Integrationszeiten quantifiziert. Das arithmetische Mittel des Positionsrauschens wurde für den Chip I7 bei einer Integrationszeit von 1 ms zu ca. 1.85 Pixel berechnet, für Chip I8 zu ca. 2.13 Pixel. Unter Verlängerung der Integrationszeit und der dadurch möglichen Verringerung der Feedback-Verstärkung reduzierte sich das Positionsrauschen bei 2 ms auf ca. 1.3 und 1.26 Pixel und bei 4 ms auf 0.74 und 0.94 Pixel.

Die Positionsdetektion der Fokalfunkte wurde anschließend eingesetzt, um Wellenfront-Aberrationen von Test-Linsen zu bestimmen. Die nominale Auflösung der Fokalfunkt-Position unter Annahme einer idealen Diskretisierung der tatsächlichen Position der Fokalfunkte in Abschnitte mit einem Abstand der Photodetektor-Streifen ermöglicht die Bestimmung der optischen Aberrationen der Wellenfront mit einem absoluten Fehler von ± 0.05 dpt (vergleiche Abschnitt 1.6.1). Unter Einstellen einer Integrationszeit von 1 ms vergrößert sich der Fehler der Positionsdetektion auf 2.13 Pixel, so dass ein absoluter Fehler in der Rekonstruktion der Wellenfront von ± 0.16 dpt

entsteht. Unter Verlängerung der Integrationsdauer auf 2 oder 4 ms resultiert ein absoluter Fehler der Wellenfront-Rekonstruktion von ± 0.12 und ± 0.10 dpt.

Abschließend wurden durch das System die optischen Aberrationen von Test-Linsen vermessen. Es handelte sich um Linsen von Sehhilfen, die über defokale und astigmatistische Aberrationen in einem Bereich von bis zu 1.25 dpt verfügten. Anhand der Vergleiche der gemessenen Aberrationen mit den für die Linsen spezifizierten Werten konnte die Funktionsfähigkeit des Systems festgestellt werden. Je nach Integrationszeit betrug der berechnete absolute Fehler der Messwerte zwischen ± 0.16 und ± 0.10 dpt, wobei die Standardabweichungen der aus 100 aufeinanderfolgenden Messungen berechneten Aberrationen im Allgemeinen unter 5% lagen.

Der tatsächlich auftretende Fehler in der Bestimmung der optischen Aberration ist aufgrund der statistischen Verteilung der Fehler der einzelnen Messwerte (Fokalkpunkt-Positionen) und deren Mittelung anhand des Least-Square-Fit-Verfahrens geringer anzunehmen. Es ist anzumerken, dass unter Verwendung einer Integrationszeit von 1 ms bei 6 von 8 vermessenen Aberrationstypen die Messung mit der Spezifikation auf mindestens 90% übereinstimmten. In 2 Ausnahmen gab es eine starke Diskrepanz der vermessenen Werte und ihrer Spezifikationen, die sich selbst unter Annahme des maximalen absoluten Fehlers des Messverfahrens nicht erklären ließ. Die Intensität der Fokalkpunkte entsprach während der Messungen $P_{FP} \approx 1 \pm 0.1 \text{ nW}$.

Eine Verbesserung der absoluten Auflösung des Systems in der Bestimmung optischer Aberrationen auf ± 0.07 dpt ist unter Verwendung von Fokalkpunktintensitäten von $J_{FP} = 200 \frac{\text{pW}}{\text{Spot}}$ und Fokalkpunkt-Durchmessern von weniger als $17 \mu\text{m}$ möglich. Nach Messungen beträgt die Ungenauigkeit der Positions-Detektion in diesem Fall ca. 30% einer Detektor-Streifen-Breite und verursacht eine absolute Ungenauigkeit der Positions-Detektion von ca. 1.3 Pixel.

Die Messungen wurden mit zwei monochromatischen Lichtquellen mit Wellenlängen von 675 und 685 nm durchgeführt. Bei Verwendung der p+/Nwell-Photodioden im HSSX beträgt die Quanteneffizienz nach Abschätzungen $\eta = 0.1$. Je nach eingesetzter Wellenlänge kann die Sensitivität der Dioden erhöht (UV) oder verringert (IR) werden, siehe Abbildung 2.4. Die im Diagramm eingetragenen Werte entstammen Simulationen und Messungen für den 0.8- μm -CMOS-Prozess der Firma AMS. Die Quanteneffizienz der verwendeten Photodioden des 0.6- μm -CMOS-Prozess wurde unter Vergleich der geometrischen Lage und der Gestalt des pn-Überganges der p-Diffusionszonen in der N-Wanne mit denen des 0.8- μm -CMOS-Prozess abgeschätzt und beträgt bei $\lambda = 680 \text{ nm}$ ca. 0.1 ± 0.02 , bei $\lambda = 800 \text{ nm}$ ca. 0.05 ± 0.01 . Unter Wahl einer kürzeren Wellenlänge mit $\lambda \approx 500 \text{ nm}$ kann die Quanteneffizienz auf 0.3 ± 0.05 erhöht werden.

Zusammenfassung und Ausblick

In dieser Arbeit wurde das System eines Hartmann-Shack-Wellenfrontsensors zur quantitativen Bestimmung der Aberrationen optischer Elemente entwickelt, aufgebaut und vermessen. Hauptbestandteil der Arbeit war die Entwicklung eines applikationsspezifischen integrierten Schaltkreises, der durch eine Kombination von Photodetektoren und elektronischer Signalverarbeitung die Positionen von Fokalfpunkten detektieren kann. Die Fokalfpunkte werden durch die Beleuchtung eines Linsenarray durch den zu vermessenden Lichtstrahl erzeugt.

Aufgrund der Anwendung des Wellenfrontsensors in einem System der Ophthalmologie zur Visualisierung der Retina wurden Spezifikationen in Bezug auf die optische Sensitivität und die Höhe der Verarbeitungsgeschwindigkeit des ASIC definiert. Die Aberrationen des zu untersuchenden Auges reduzieren die Auflösung des Visualisierungsverfahrens. Durch Messen der optischen Aberrationen des Auges und Veränderung der Phaseninformation des Mess-Strahls des Visualisierungs-Systems durch einen aktiven Spiegel ist es möglich, die Auflösung des Verfahrens zu verbessern.

Aufgrund der geringen Intensität des an der Retina reflektierten Messstrahls resultiert eine Lichtleistung der Fokalfpunkte des Linsenarray von ca. 200 pW bis 1 nW. Zusammen mit der Forderung nach einer hohen Messgeschwindigkeit des Wellenfrontsensors von 1 kHz aufgrund der inhärenten Bewegungen des Auges entstanden Spezifikationen, deren Erfüllung Maßgabe in der Entwicklung des ASIC war.

Der ASIC verwendet zur Bestimmung der Fokalfpunkt-Positionen die im 0.6- μm -CMOS-Prozess realisierbaren p+/nwell-Photodioden, die eine Quanteneffizienz von 0.15 bei einer Wellenlänge von 680 nm und von 0.1 bei einer Wellenlänge von 780 nm besitzen. Die Positionsbestimmung erfolgt durch die Anwendung einer Winner-Take-All-Schaltung, die unter verschiedenen Photostromen von Photodioden-Streifen den stärksten Photostrom detektiert. Die Photodioden-Streifen, die in X- und Y-Richtung in einem Abstand von 17 μm angeordnet sind, werden durch die Fokalfpunkte beleuchtet. Durch analoge Signalverarbeitung wird der stärkste Photostrom erkannt und die Position quantisiert. Die generierten Daten werden durch ein externes System, bestehend aus einem Dual-Port-RAM und einem FPGA, an einen PC übertragen, der durch ein Least-Square-Fit-Verfahren der Daten die Koeffizienten von Zernike-Polynomen berechnet.

Innerhalb der Arbeit wurden zwei Teststudien des Wellenfrontsensor-ASIC mit einer geringen Anzahl von Clustern entworfen. Aus den Erkenntnissen der Untersuchungen entstand der letzte Prototyp *HSSX*, der mit einer Anzahl von 16×16 Clustern die Position von 256 Fokalfpunkten des im Hartmann-Shack-Sensors eingesetzten Linsenarray mit einer nominalen Auflösung von 17 μm in einem dynamischen Bereich von 340 μm vermessen kann. Das Verhalten des *HSSX* wurde in Messungen überprüft.

Unter den insgesamt 180 prozessierten Chips wurden Exemplare ausgewählt, bei denen eine gute Übereinstimmung der Transistor-Charakteristika von identischen, benachbarten Strukturen festgestellt werden konnte. Mit zwei Chips wurde ein Hartmann-Shack-Sensor aufgebaut und die optischen Aberrationen von Testlinsen vermessen.

Bei einer Integrationszeit von 1 ms war es möglich, die Position der Fokalfunkte mit einer absoluten Genauigkeit von 3 Streifen zu detektieren, entsprechend $51\mu\text{m}$. In einem dynamischen Bereich von 19 Streifen entspricht dies einer Genauigkeit von 15%. Umgerechnet auf den dynamischen Bereich der optischen Aberrationen von ± 1 dpt erfolgt eine absolute Genauigkeit der Messung von ± 0.15 dpt. Bei Verlängerung der Integrationszeit auf 4 ms war es möglich die absolute Genauigkeit auf ± 0.10 dpt zu verbessern. Die Verbesserung der Genauigkeit bei längeren Integrationszeiten resultiert aus der geringeren Stärke der erforderlichen Rückkopplung der Photoströme und dem dadurch verringerten Rauschen der Positionsdetektion.

Durch Messungen konnte festgestellt werden, dass unter Verwendung von Fokalfunkten mit kleinerem Durchmesser als die von dem Linsenarray erzeugten $200\mu\text{m}$ eine weitere Verbesserung der Genauigkeit erzielt werden könnte. Bei Messung eines Fokalfunktes mit einer Lichtleistung von 200 pW, einem maximalen Durchmesser von $17\mu\text{m}$ und mit einer Integrationszeit von 1 ms war eine Genauigkeit der Positionsdetektion von 1.3 Streifen möglich. Unter Annahme dieser Detektionsgenauigkeit im Hartmann-Shack-Verfahren würde eine Genauigkeit in der Bestimmung der optischen Aberrationen von ± 0.07 dpt erreicht werden können.

Verbesserungsmöglichkeiten des *HSSX* zur Erhöhung der optischen Sensitivität und der Detektionsgenauigkeit liegen in der Anwendung anderer Photodetektoren mit höherer effektiver Quanteneffizienz und der Reduzierung der Fokalfunkt-Durchmesser, um höhere Fokalfunktintensitäten und dadurch höhere Photoströme bei gleicher Lichtleistung zu erhalten.

Das System der Aberrationskompensation zur Auflösungsverbesserung des Visualisierungs-Verfahrens der Retina konnte im Zeitraum des Projektes nicht aufgebaut werden, da sich die Herstellung des erforderlichen aktiven Spiegels verzögerte. Die angeführten Messungen ergaben eine Funktionsfähigkeit eines Laborsystems des Wellenfrontensors, wobei während der Entwicklung des ASIC die charakteristischen Eigenschaften des Systems in Bezug auf optische Sensitivität, Auflösung und Verarbeitungsgeschwindigkeit in Hinblick auf den Einsatz in dem Visualisierungsverfahren optimiert wurden.

Die Anwendung des in diesem Projekt entwickelten Wellenfrontensors als Aberrationsdetektor in einem System zur Auflösungsverbesserung eines Visualisierungsverfahrens der Retina ist daher Gegenstand der weiteren, zukünftigen Forschungen.

Anhang A

Least-Square-Fit

Nach den Ausführungen im Kapitel 1.5.2 werden die Zernike-Koeffizienten zur mathematischen Modellierung der Wellenfrontaberrationen durch einen Least-Square-Fit der Messpunkte angenähert. Die Differenz zwischen dem Messwert und der Annäherung über die Summen-Funktion wird für jede Mess-Stelle quadriert. Die Koeffizienten der Polynome werden nun derart berechnet, daß die Summe aller Differenze-Quadrate einen minimalen Betrag darstellt.

Unter Berücksichtigung der partiellen Ableitungen der Zernike-Polynome nach x, y ergeben sich die folgenden Beziehungen (N sei die Anzahl der Messpunkte):

$$\begin{aligned} S_x &= \sum_{i=1}^N \left(\frac{\partial}{\partial x} W(x, y) \Big|_{x=x_i, y=y_i} - \frac{\Delta x_i}{f} \right)^2 \\ S_y &= \sum_{i=1}^N \left(\frac{\partial}{\partial y} W(x, y) \Big|_{x=x_i, y=y_i} - \frac{\Delta y_i}{f} \right)^2 \\ W(x, y) &= \sum_{j=1}^{14} C_j \cdot Z_j(x, y) \end{aligned} \quad (\text{A.1})$$

Um die Summen S_x und S_y zu minimieren, werden die partiellen Ableitungen von S_x und S_y nach C_k Null gesetzt:

$$\begin{aligned} \frac{\partial S_x}{\partial C_k} &= \frac{\partial}{\partial C_k} \sum_{i=1}^N \left(\frac{\partial}{\partial x} \sum_{j=1}^{14} C_j \cdot Z_j(x, y) - \frac{\Delta x_i}{f} \right)^2 = 0 \quad \forall k \in \{1, 2, \dots, 14\} \\ \frac{\partial S_y}{\partial C_k} &= \frac{\partial}{\partial C_k} \sum_{i=1}^N \left(\frac{\partial}{\partial y} \sum_{j=1}^{14} C_j \cdot Z_j(x, y) - \frac{\Delta y_i}{f} \right)^2 = 0 \quad \forall k \in \{1, 2, \dots, 14\} \end{aligned} \quad (\text{A.2})$$

Daraus entstehen die folgenden Beziehungen:

$$\begin{aligned} \sum_{i=1}^N \frac{\partial}{\partial x} \sum_{j=1}^{14} C_j \cdot Z_j(x, y) \cdot \frac{\partial}{\partial x} Z_k(x, y) &= \sum_{i=1}^N \frac{\Delta x_i}{f} \cdot \frac{\partial}{\partial x} Z_k(x, y) \quad \forall k \in \{1, 2, \dots, 14\} \\ \sum_{i=1}^N \frac{\partial}{\partial y} \sum_{j=1}^{14} C_j \cdot Z_j(x, y) \cdot \frac{\partial}{\partial y} Z_k(x, y) &= \sum_{i=1}^N \frac{\Delta y_i}{f} \cdot \frac{\partial}{\partial y} Z_k(x, y) \quad \forall k \in \{1, 2, \dots, 14\} \end{aligned} \quad (\text{A.3})$$

Die Lösung der oben stehenden Gleichungen kann vereinfacht werden, wenn für die partiellen Ableitungen der Zernike-Polynome nach x, y Ersatz-Polynom-Funktionen verwendet werden, die sich bezüglich ihrer Polynome orthogonal verhalten. Es seien

$$\begin{aligned} \sum_{l=1}^{10} r_l \cdot P_l(x, y) &= \frac{\partial}{\partial x} \sum_{j=1}^{14} C_j \cdot Z_j(x, y) \\ \sum_{l=1}^{10} s_l \cdot P_l(x, y) &= \frac{\partial}{\partial y} \sum_{j=1}^{14} C_j \cdot Z_j(x, y) \end{aligned} \quad (\text{A.4})$$

Die Ersatzpolynome P_i werden durch das Gram-Schmidt Orthogonalisierungsverfahren nach [Malacara 92] gebildet und lauten:

P_1	1
P_2	x
P_3	y
P_4	xy
P_5	$x^2 - y^2$
P_6	$x^2 + y^2 - \frac{2 \sum_i x_i^2}{N}$
P_7	$x^3 - x \cdot \frac{2 \sum_i x_i^4}{\sum_i x_i^2}$
P_8	$-y^3 + y \cdot \frac{2 \sum_i x_i^4}{\sum_i x_i^2}$
P_9	$xy^2 - x \cdot \frac{2 \sum_i x_i^2 y_i^2}{\sum_i x_i^2}$
P_{10}	$x^2 y - y \cdot \frac{2 \sum_i x_i^2 y_i^2}{\sum_i x_i^2}$

Tabelle A.1: Die orthogonalen Ersatzpolynome

Nach Einsetzen von A.4 in A.3 und unter Anwendung des Orthogonalitätsprinzips

$$P_l(x, y) \cdot P_m(x, y) = 0 \quad \forall l \neq m \quad (\text{A.5})$$

ergeben sich für die Koeffizienten der Ersatz-Polynom-Funktionen die folgenden Relationen:

$$\begin{aligned} r_l &= \sum_{i=1}^N \frac{\Delta x_i}{f} P_l(x_i, y_i) / \sum_{i=1}^N P_l^2(x_i, y_i) \\ s_l &= \sum_{i=1}^N \frac{\Delta y_i}{f} P_l(x_i, y_i) / \sum_{i=1}^N P_l^2(x_i, y_i) \quad \forall l \in \{1, 2, \dots, 10\} \end{aligned} \quad (\text{A.6})$$

Durch Koeffizientenvergleich der Terme von $x^i \cdot y^j$, $i, j \in \{0, 1, 2, 3\}$, $i \cdot j < 3$ in A.3 wird ein lineares Gleichungssystem aus 20 Gleichungen gebildet. Davon sind 5 Gleichungen voneinander linear abhängig und 2 Gleichungen erzeugen eine Überbestimmung. Nach Eliminierung der fünf linear abhängigen und einer der beiden überbestimmenden Gleichungen entsteht ein für alle C_i , $i \in \{1, 2, \dots, 14\}$ lösbares System aus 14 Gleichungen; es verfügt über die folgenden Lösungen:

C_1	$r_1 + (\frac{1}{3} - \alpha)r_6$
C_2	$s_1 + \frac{1}{12}r_4 - \frac{1}{6}s_5 + (\frac{1}{6} - \alpha)s_6$
C_3	$\frac{1}{2}r_3 + (\frac{1}{2}\beta - \frac{3}{16})r_8 + (\frac{1}{16} - \frac{1}{2}\tau)r_{10}$
C_4	$\frac{1}{8}r_2 + \frac{1}{8}s_3 + (\frac{3}{64} - \frac{1}{8}\beta)r_7 + (\frac{1}{32} - \frac{1}{8}\tau)r_9 + (\frac{1}{8}\beta - \frac{3}{64})s_8 - \frac{1}{8}s_{10}\tau$
C_5	$-\frac{1}{4}r_2 + \frac{1}{4}s_3 + (\frac{1}{4}\beta - \frac{3}{32})r_7 + (\frac{1}{4}\beta - \frac{3}{32})s_8 + \frac{1}{4}r_9\tau - \frac{1}{4}s_{10}\tau$
C_6	$-\frac{1}{3}r_5 + \frac{1}{6}r_6$
C_7	$\frac{1}{6}r_6$
C_8	$\frac{1}{24}r_4 - \frac{1}{12}s_5 + \frac{1}{12}s_6$
C_9	$-\frac{1}{8}r_4 - \frac{1}{12}s_5 + \frac{1}{12}s_6$
C_{10}	$-\frac{1}{8}r_8 - \frac{1}{24}r_{10}$
C_{11}	$-\frac{1}{16}r_8 + \frac{1}{48}r_{10}$
C_{12}	$\frac{1}{64}r_7 - \frac{1}{64}s_8 + \frac{1}{96}r_9$
C_{13}	$-\frac{1}{32}r_7 - \frac{1}{32}s_8$
C_{14}	$\frac{1}{32}r_7 - \frac{1}{32}s_8 - \frac{1}{16}r_9$

Tabelle A.2: Umrechnung der Ersatz-Koeffizienten in Zernike-Koeffizienten

Die Werte von α , β und τ lauten:

$$\begin{aligned}
 \alpha &= 2 \frac{\sum_{i=1}^N x_i}{N} \\
 \beta &= \frac{\sum_{i=1}^N x_i^4}{\sum_{i=1}^N x_i^2} \\
 \tau &= \frac{\sum_{i=1}^N x_i^2 y_i^2}{\sum_{i=1}^N x_i^2}
 \end{aligned} \tag{A.7}$$

Anhang B

Verwendete Variablen

A_S	Sphärische Aberration in [dpt]
A_Z	Astigmatische Aberration in [dpt]
α	Winkel einer Wellenfront-Verkippung
β	Verstärkungsfaktor
β	Rotation des Auges
C	Zernike-Koeffizient in [μm]
C^*	Zernike-Koeffizient in [$\frac{1}{m}$]
ΔC	Koeffizienten-Differenz
c_j, C_j	Sperrschicht-Kapazität
C_{ox}	Oxyd-Kapazität pro Fläche
d	Durchmesser eines Fokal-Punktes
D	Durchmesser der Apertur
D_A	Aberration in [dpt]
D_{FP}	Durchmesser des Fokal-Punktes
D_L	Durchmesser einer Linse/Pupille
D_n	Diffusionskonstante von Elektronen
D_p	Diffusionskonstante von Defektelektronen
Δf	Frequenz-Bandbreite der Anwendung
e	Elementarladung
E	Elektrische Feldstärke
E	Erwartungswert
ϵ_{si}	Dielektrizitätskonstante von Silizium
f	Fokal-Länge des Linsenarrays
f_L	Fokal-Länge des Linsenarrays
f_D	Fokal-Länge einer Fokussierlinse
Φ_f	Oberflächenpotential in starker Inversion
g_m	Übertragungsleitwert / Transkonduktanz
γ	Faktor des Body-Effektes
h	Plank'sches Wirkungsquantum
i	Index
i_s, I_s	Sperr-Sättigungsstrom
i_n, I_n	Rausch-Strom

I_D	Drain-Strom
$I(x)$	Strahlungs-Intensität an der Stelle x
j	Index
$J_1(z)$	Bessel-Funktion erster Gattung erster Ordnung
J_{ph}	Intensität elektromagnetischer Strahlung in $[W/m^2]$
J_{FP}	Strahlungs-Intensität eines Fokal-Punktes
λ	Wellenlänge
k	Ausbreitungsvektor; Boltzmannkonstante
KF	Prozess-Parameter zur Modellierung des $1/f$ -Rauschens
L	Kanallänge
L_n	Diffusionslänge von Elektronen
L_p	Diffusionslänge von Defektelektronen
λ	Parameter zur Modellierung der Kanallängen-Modulation
λ	Wellenlänge elektromagnetischer Strahlung
n	Index
n_i	Intrinsische Ladungsträgerkonzentration ($n_i = 1.14 \cdot 10^{10} \text{ cm}^{-3}$)
N_A	Konzentration der Akzeptor-Atome
N_D	Konzentration der Donator-Atome
ν	Frequenz elektromagnetischer Strahlung in $[Hz]$
P_{ph}	Leistung elektromagnetischer Strahlung in $[W]$
P_{FP}	Strahlungs-Leistung eines Fokal-Punktes
ϕ	Phaseninformation
Φ	Phaseninformation als $f(x,y)$
r_A	Radius
r_1	Radius des Auges
r_2	Radius der Hornhautkrümmung
S	Verhältnis W/L
σ	Standardabweichung
t	Zeit
T	Temperatur
Θ	Winkel des Astigmatismus in $[^\circ]$
V_T	Temperatur-Spannung kT/e
V_{BS}	Bulk-Source-Spannung
V_{GS}	Gate-Source-Spannung
V_{DS}	Drain-Source-Spannung
W	Wellenfront oder Phasendifferenz, Kanalbreite
x	Koordinate, einheitenlos oder in $[m]$
Δx	Verschiebung eines Fokal-Punktes
Δx_d	Detektor-Auflösung
Δx_{max}	Maximal mögliche Verschiebung
y	Koordinate, einheitenlos oder in $[m]$
Δy	Translatorische Verschiebung des Auges
z	Tiefe im Halbleiter / Strecke entlang der optischen Achse
Z_i	Zernike-Polynom

Literaturverzeichnis

- [Allen 87] P. Allen, "CMOS analog circuit design", Oxford University Press, 1987, S. 66
- [Allen 87b] [Allen 87], S. 173
- [Allen 87c] [Allen 87], S. 99
- [Allen 87d] [Allen 87], S. 206
- [Allen 87e] [Allen 87], S. 234
- [AMS 0.8] Austria Mikro Systeme International AG, "0.8 μm CMOS Process Parameters"
- [AMS 0.6 I] Austria Mikro Systeme International AG, "0.6 μm CMOS Joint Group Process Parameters"
- [AMS 0.6 II] Austria Mikro Systeme International AG, "0.6 μm CMOS Design Rules"
- [AMS 0.6 III] Austria Mikro Systeme International AG, "0.6-Micron Standard Cell Data-book"
- [AMS 0.6 IV] Austria Mikro Systeme International AG, "ESD Design Rules"
- [Artal 88] P. Artal et al., "Retrieval of wave aberrations of human eyes from actual point-spread-data", Journal of the Optical Society of America A, Vol. 5 , pp. 1201-1206, 1988
- [Artal 94] P. Artal et al., "Odd aberrations and double-pass measurements of retinal image quality", Journal of the Optical Society of America A, Vol. 12 , pp. 195-201, 1995
- [Bergmann 92] Bergmann, Schäfer: "Festkörper", Gruyter, 1992
- [Born 65] Born, Wolf, "Principles of Optic", Pergamon, 1965, S. 464 ff
- [Campbell 66] Campbell and Gubisch, "Optical quality of the human eye", J. Physiol. London 186, 558, 1966
- [Carpenter 88] C. Carpenter, "Movements of the Eyes", Pion Limited, 1988
- [Decker 98] S. Decker et al., "A 256×256 CMOS Imaging Array with Wide Dynamic Range Pixels and Column-Parallel Digital Output", Journal of Solid State Circuits, Vol. 33, No. 12, p. 2081, 1998
- [Dreher 89] Dreher et al., "Active optical depth resolution improvement of the laser tomographic scanner", Journal of Applied Optics, Vol 28 No 4, pp 804, 1989

- [Geary 95] J. Geary, "Introduction to wavefront sensors", SPIE Press, Vol. TT18, pp. 89, 1995
- [Geiger 90] R. Geiger, "VLSI Design Techniques for Analog and Digital Circuits", McGraw-Hill, 1990
- [Gregorian 86] R. Gregorian, "Analog MOS integrated circuits", Wiley and Sons, 1986
- [Grey 84] R. Gray, "Analog Integrated Circuits, 2nd Edition", Wiley and Sons, 1984
- [Grey 84b] [Grey 84], S. 689
- [Jones 87] K. Jones, "Optoelektronik", Harper And Row, 1987, S. 234
- [Langeheine 98] J. Langeheine, "Entwurf und Test eines optisch konfigurierbaren Widerstandsnetzwerkes", Diplomarbeit, IHEP, 1998
- [Liang 97] Liang et al., "Supernormal vision and high-resolution retinal imaging through adaptive optics", Journal of the Optical Society of America A, Vol. 14, pp 2884, 1997
- [Liang 94] J. Liang et al., "Objective measurements of wave aberrations of the human eye with the use of a Hartmann-Shack wave-front sensor", Journal of the Optical Society of America A, Vol. 11, pp. 1949-1957, 1994
- [Loose 96] Markus Loose, "Layout und Test eines Systems adaptiver Photorezeptoren in analoger CMOS-Technologie", Diplomarbeit, IHEP, 1996
- [Maxim] Maxim Incorporation, "Datasheets MAX510, MAX5120"
- [Malacara 92] D. Malacara, "Optical shop testing", Wiley and Sons, 1992
- [Merkle 88] F. Merkle et al., "Aktive und adaptive Optik in der Astronomie", Physikalische Blätter 44 Nr. 12, pp. 439-446, 1988
- [Wolf 71] H.F. Wolf, "Semiconductors", Wiley and Sons, 1971
- [Microsoft 93] Microsoft Corporation, "Extendet Capabilities Port: Specifications, Revision 1.06", Microsoft, 1993
- [Schneider 98] Peter Schneider, "Simulation und Visualisierung elektrischer und optischer Eigenschaften von Halbleiterbauelementen", Diplomarbeit, IHEP, 1998
- [Seiler 98] T. Seiler, Presseerklärung, 96. Tagung DOAG, 19. - 22. September 1998, Berlin
- [Sze 85] Sze, "Semiconductor devices - physics and technology", Wiley and Sons, 1985
- [Teodori 97] Marcus Teodori, "Design und Messung optoelektronischer Strukturen in CMOS Technologie", Diplomarbeit, IHEP, 1997
- [Vittoz 77] E. Vittoz et al., "CMOS Analog integrated circuits based on weak inversion operation", Journal of Solid State Circuits, Vol. SC-12, No. 3, S. 224 ff, 1977
- [Vittoz 96] E. Vittoz, "Bioinspired Signal Processing", Workshop
- [Zernike 34] F. Zernike, "Beugungstheorie des Schneidenverfahrens und seine verbesserte Form der Phasenkontrastmethode", Physica, 1, 689-704, 1934

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mir die Möglichkeit zur Anfertigung dieser Arbeit gegeben haben. Mein persönlicher Dank gilt den Mitarbeitern und Kollegen des ASIC-Labors der Universität Heidelberg, die durch ihre herzliche und freundliche Art ein besonders angenehmes Arbeitsklima entstehen ließen.

An erster Stelle möchte ich mich bei meiner Mutter und im Gedenken bei meinem Vater für ihre ständige Unterstützung meines Lebensweges und die Ermöglichung meiner Ausbildung bedanken.

Desweiteren gilt mein Dank den folgenden Personen:

Herrn Prof. Josef Bille für die Übergabe der Bearbeitung dieses sehr interessanten Projektes, für die Betreuung der Arbeit während der letzten 3 Jahre und für die Anfertigung des Zweitgutachtens,

Herrn Prof. Reinhard Männer für die Ermöglichung der Anfertigung der Arbeit im Rahmen des Kooperationsabkommens der Universitäten Mannheim und Heidelberg und für das Anfertigen des Erstgutachtens,

Herrn Prof. Karlheinz Meier für die freundliche Zusammenarbeit und vor allem für seinen extremen Einsatz für die Forschung und Lehre, insbesondere die hervorragende Infrastruktur des ASIC-Labors,

Herrn Michael Keller, dem ich an dieser Stelle meinen tiefen Respekt und meinen Dank für die Hilfestellungen in jedem Bereich des Lebens und für den unermüdlichen Einsatz zur Aufrechterhaltung der Funktion der Datenverarbeitungssysteme aussprechen möchte,

Herrn Jörg Langeheine für die angenehme Arbeitsatmosphäre und für die interessanten Gespräche,

Herrn Johannes Schemmel und Herrn Markus Loose für ihre spontanen Hilfestellungen in allen Fragen des wissenschaftlichen Lebens und für die sehr interessanten Diskussionen,

Herrn Dr. Martin Feuerstack-Raible für die Organisation des ASIC-Labors und der Meetings und für die freundliche Unterstützung,

Herrn Peter Schneider für die seinerzeits sehr angenehme Zusammenarbeit

und Herrn Thorsten Maucher für seine freundliche und zuvorkommende Art und die problemlose Absprache in der Verwendung der Labtops.

Ebenfalls bedanken möchte ich mich bei den Kollegen und Mitarbeitern des Instituts für Angewandte Physik, besonders bei

Herrn Dr. Stephan Gölz für seine Bemühungen um die Forschung und für die sehr gute fachliche Beratung in vielen Fragen,

Herrn Ingo Müller-Vogt für die effiziente und strukturierte Teamarbeit im Sommer 1998 und für die jederzeit freundliche Art

und bei den Herren Alexander Türpitz, Ralf Kessler und Uli von Pape für ihre Hilfe und Unterstützung während der letzten Jahre.

Den Damen und Herren Irmi Spinnler, Elke Grimm, Ellen Schmidt, Lothar Vörg, Dr. Oskar Braun, Reiner Knebel und allen anderen Mitarbeitern des Instituts für Hochenergiephysik möchte ich ebenfalls sehr herzlich für ihren Einsatz und die Zusammenarbeit danken.

Mein ganz besonderer Dank gilt meinen Freundinnen und Freunden, die mich durch die Höhen und Tiefen der letzten Jahre begleitet haben und auf die ich mich jederzeit verlassen konnte.

Ich versichere, dass ich die vorliegende Arbeit ohne unzulässige Hilfe und nur unter Verwendung der angegebenen Mitteln angefertigt habe.

Heidelberg, den 14.5.1999

